

# INTERDISCIPLINARY DESCRIPTION OF COMPLEX SYSTEMS

## Scientific Journal

---

<i>I. Shimada and T. Koyama</i>	342	A Theory for Complex Systems Social Change: An Application of a General 'Criticality' Model
<i>H. Hashim and S. Radhakrishna</i>	354	The Concept of Free Will as an Infinite Metatheoretic Recursion
<i>A.W. Gluszek and F. Kluza</i>	367	The Universe Is Like a Hollowed Sphere. The Wave Concept of Time
<i>G.P. Sarma</i>	373	The Art of Memory and the Growth of the Scientific Method
<i>G.P. Sarma</i>	397	Reconsidering Written Language
<i>M. Givel</i>	405	The Paradox of Explosive and Gradual Policy Change in Political Revolutionary Times
<i>K.Dumičić, V. Bučevska and Emina Rešić</i>	420	Recent Impacts of Selected Development Indicators on Unemployment Rate: Focusing the SEE Countries
<i>H. Miloloža</i>	434	Distance Factors and Croatian Export Obstacles in EU15: CAGE Approach
<i>Y. Addisu</i>	450	The Effect of Shocks: An Empirical Analysis of Ethiopia
<i>G. Rozing, V. Alar and V. Marušić</i>	461	Study of Stainless Steel Resistance in Conditions of Tribocorrosion Wear
<i>B. Runje, E. Krstić Vukelja and A. Horvatić</i>	472	Comparison of Different Simulations Methods in Case of Service-providing Companies
<i>Ž. Alar, D. Mandić, A. Dugorepec and M. Sakoman</i>	479	Application of Instrumented Charpy Method in Characterisation of Materials

## Scientific Journal

# INTERDISCIPLINARY DESCRIPTION OF COMPLEX SYSTEMS

INDECS, volume 13, issue 3, pages 342-487, year 2015

Published 28 July 2015 in Zagreb, Croatia

Released online 28 July 2015

## Editors

*Josip Stepanić, Editor-in-Chief*

editor@indec.s.eu; Faculty of Mechanical Engineering & Naval Architecture, I. Lučića 5, HR – 10000, Zagreb, Croatia

*Josip Kasač, Assistant Editor*

josip.kasac@indec.s.eu; Faculty of Mechanical Engineering & Naval Architecture, I. Lučića 5, HR – 10000, Zagreb, Croatia

*Mirjana Pejić Bach, Assistant Editor*

mirjana.pejic.bach@indec.s.eu; Faculty of Economics and Business, Trg J.F. Kennedyja 6, HR – 10000, Zagreb, Croatia

## Advisory Board

*Vjekoslav Afrić*, University of Zagreb, Zagreb (HR)

*Aleksa Bjeliš*, University of Zagreb, Zagreb (HR)

*Predrag Ćosić*, University of Zagreb, Zagreb (HR)

*Marek Frankowicz*, Jagiellonian University, Krakow (PL)

*Katalin Martinás*, Eötvös Loránd University, Budapest (HU)

*Gyula Mester*, University of Szeged, Szeged (HU)

*Dietmar Meyer*, Budapest University of Technology and Economy, Budapest (HU)

*Sibila Petlevski*, University of Zagreb, Zagreb (HR)

*Wei-bin Zhang*, Ritsumeikan Asia Pacific University, Beppu (JP)

## Editorial Board

*Serghey A. Amelkin*, Program Systems Institute, Pereslavl-Zalesskij (RU)

*Željko Božić*, University of Zagreb, Zagreb (HR)

*Nikša Dubreta*, University of Zagreb, Zagreb (HR)

*Robert Fabac*, University of Zagreb, Varaždin (HR)

*Mirta Galešić*, Max Planck Institute for Human Development, Berlin (DE)

*Erik W. Johnston*, University of Arizona, Ann Arbor (US)

*Urban Kordeš*, University of Ljubljana, Ljubljana (SI)

*Dean Korošak*, University of Maribor, Maribor (SI)

*Anita Lee-Post*, University of Kentucky, Lexington (US)

*Olga Markič*, University of Ljubljana, Ljubljana (SI)

*Damir Pajić*, University of Zagreb, Zagreb (HR)

*Petra Rodik*, University of Zagreb, Zagreb (HR)

*Biserka Runje*, University of Zagreb, Zagreb (HR)

*Armano Srblijinović*, Institute for R&D of Defence Systems, Zagreb (HR)

*Karin Šerman*, University of Zagreb, Zagreb (HR)

---

Published by *Croatian Interdisciplinary Society* (<http://www.idd.hr>) four times per year as printed (ISSN 1334-4684) and online (ISSN 1334-4676) edition. Printed by *Redak d.o.o.* (HR) in 50 pieces. Online edition, <http://indec.s.eu>, contains freely available full texts of published articles.

Journal INDECS is financially supported by Croatian Ministry of Science, Education and Sport.

Full content of the journal INDECS is included in the EconLit, EBSCO, Ulrich's and Index Copernicus.

INDECS publishes original, peer-reviewed, scientific contributions prepared as reviews, regular articles and conference papers, brief and preliminary reports and comments to published articles.

The accessibility of all URLs in the texts was checked one week before the publishing date.

## TABLE OF CONTENTS

<i>Ichiro Shimada and Tomio Koyama</i>	342	A Theory for Complex Systems Social Change: An Application of a General ‘Criticality’ Model
<i>Hashim Hanaan and Srikanth Radhakrishna</i>	354	The Concept of Free Will as an Infinite Metatheoretic Recursion
<i>Andrzej W. Głuszak and Franciszek Kluza</i>	367	The Universe is Like a Hollowed Sphere. The Wave Concept of Time
<i>Gopal P. Sarma</i>	373	The Art of Memory and the Growth of the Scientific Method
<i>Gopal P. Sarma</i>	397	Reconsidering Written Language
<i>Michael Givel</i>	405	The Paradox of Explosive and Gradual Policy Change in Political Revolutionary Times
<i>Ksenija Dumicic, Vesna Bucevska and Emina Resic</i>	420	Recent Impacts of Selected Development Indicators on Unemployment Rate: Focusing the SEE Countries
<i>Helena Miloloža</i>	434	Distance Factors and Croatian Export Obstacles in the EU15: CAGE Approach
<i>Yilebes Addisu Damtie</i>	450	The Effect of Shocks: An Empirical Analysis of Ethiopia
<i>Goran Rozing, Vesna Alar and Vlatko Marušić</i>	461	Study of Stainless Steel Resistance in Conditions of Tribocorrosion Wear
<i>Biserka Runje, Elizabeta Krstić Vukelja and Amalija Horvatić</i>	472	Comparison of Different Simulations Methods in Case of Service-Providing Companies
<i>Željko Alar, Davor Mandić, Andrija Dugorepec and Matija Sakoman</i>	479	Application of Instrumented Charpy Method in Characterisation of Materials
<i>Abstracts in Croatian</i>	ii	

# A THEORY FOR COMPLEX SYSTEM'S SOCIAL CHANGE: AN APPLICATION OF A GENERAL 'CRITICALITY' MODEL

Ichiro Shimada<sup>1,\*</sup> and Tomio Koyama<sup>2</sup>

<sup>1</sup>Department of Developmental Biology and Neurosciences, Graduate School of Life Sciences  
– Tohoku University

Sendai, Japan

<sup>2</sup>Institute for Materials Research – Tohoku University

Sendai, Japan

DOI: 10.7906/indecs.13.3.1

Regular article

*Received:* 3 June 2015.

*Accepted:* 23 June 2015.

## ABSTRACT

Within the developed nations deterioration in the basis of society, as dramatically demonstrated by the Lehman collapse, has reached extreme levels, and currently the formation of pro-change agents is approaching a decisive stage. Here, we will construct a complex systems 'criticality' model, apply it to social change, and examine its reliability and validity. The model derived a power law distribution of the output of social change. The validity of the model was verified by examining vote shares of parties in Japan. Based on the results of this examination, we propose a new quantitative strategy "information entropy enhancement" for social change.

## KEY WORDS

complex systems social change, criticality, power-law distribution, information entropy enhancement strategy, vote share

## CLASSIFICATION

JEL: C65, P16

PACS: 89.75.-k, 89.65.Ef

\*Corresponding author, *η*: i-shimada@jcom.home.ne.jp; +81 022 277 4760;  
Department of Developmental Biology and Neurosciences, Graduate School of Life Sciences,  
Tohoku University, Katahira 2-1-1, Aoba-Ku, Sendai 980-8578, Japan

## INTRODUCTION

Complex societies, such as the developed capitalist country of Japan, are networks of complex systems that evolve around the focal point of struggle between pro-change and pro-establishment powers. The pro-change forces are in turn composed of networks of complex systems having diverse connections (information, needs, movements, cooperation/collaboration etc.) between various individuals and groups, in order to break free of existing systems and effect a shift to a new system. Social change manifests in the 'critical state' of mutual antagonism and conflict between the pro-change and pro-establishment powers, while the dimensions of the change (quantitative and qualitative) are defined by the network pattern of the pro-change power. Simultaneously, the heretofore dominant pro-establishment power is relegated to a secondary role. In this article, we will generalize the quantitative theory of adaptive animal behavior [1-3] into a 'criticality' model of complex systems (Appendix), apply it to the quantitative aspects of social change, construct a model for complex systems social change, and examine the model's effectiveness.

## A COMPLEX SYSTEMS SOCIAL CHANGE MODEL

The criticality model of complex systems can be applied to the dynamics of complex systems social change. Here, the pro-change power of the complex system society is considered to be the confrontation power, constructed from complex system networks of diverse connections between various individuals and groups for the purpose of breaking free of the existing systems of the complex system society and effecting a shift to a new system. We will define the number of diverse and active individuals and organizations that structure the pro-change power's network and are related to the output  $R$  as  $N$ . Following the criticality model, the number of network patterns, for example, an information exchange quantity (information sorts  $\times$  frequency) is hypothesized as  $I(\alpha N) \sim \exp(\alpha N)$  (see expression (1) in Appendix<sup>1</sup>).

It does not touch upon the details of mutual interactions between the various elements of the two conflicting powers. Hereafter, figures in parentheses indicate equations in Appendix. The existence probability of the number of network patterns of the pro-change power is given by  $P(\alpha N)$  (2), while the existence (resistance) probability of the pro-establishment power acting upon the pro-change network is given by  $Q(\beta N)$  (3). Assuming the social change output  $R$  is inversely proportional to the resistance probability  $Q(\beta N)$  (4), we can use it to obtain the probability density function  $\phi(R)$  (6) through variable transformation by  $R$  of the existence probability of the pro-change power,  $P(\alpha N)$ . Approximating the information entropy of the network of a social pro-change power  $N$  using  $H_M$  (9) leads to  $\alpha_M$  (12), which incorporates the information effect in  $\alpha$ . The criticality model of complex systems leads logically to the power law  $\Phi(R) \sim R^{-D_M}$  ( $D_M = \alpha_M/\beta$ ) for the social change output  $R$ . It should be noted that the size (strength) of the pro-change power defined by  $\alpha_M$  and  $N_C$  (largest of  $N$ ) does not directly correspond with the individual change outputs of  $R$ . Even if the pro-change power is constant, the change output  $R$  will be unstable and fluctuate greatly, following a power law distribution. The parameter of existence probability for the pro-change power is  $\alpha_M$ ; it is opposed to the pro-establishment power's resistance probability parameter  $\beta$  in contributions to the power exponent  $D_M$ . Consequently,  $D_M$  reflects relatively each agonistic probability, and therefore, real power relationship of both powers. Thus, it is difficult to grasp the agonistic structure directly from the output  $R$ , which expresses complex factors, but it can be captured via the power exponent  $D_M$ .

In order to increase the probability of true social change, the agonistic structure expressed by the power exponent must be greatly altered. To do this, the power exponent  $D_M$  should be

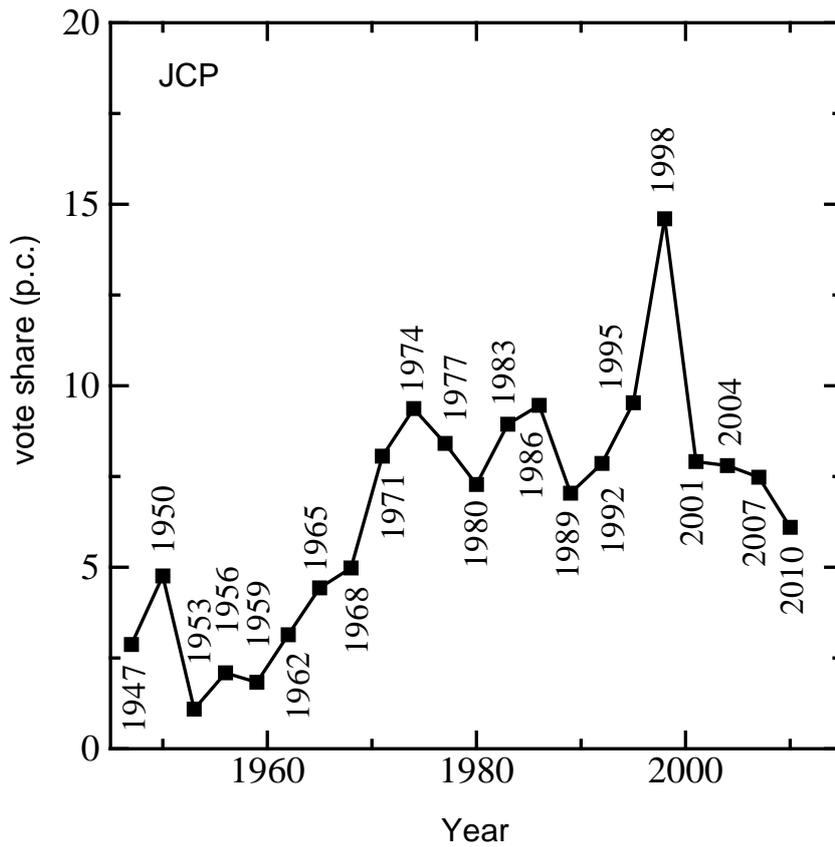
decreased so that the change output should dramatically increase. In other words,  $\alpha_M \sim \alpha_0 \exp(\Delta H)$  must be small (decrease  $\Delta H$ ), while  $\beta$  must be large, and the output  $R$  should be greatly increased. In particular,  $\Delta H$  is very effective as it contributes exponentially to  $D_M$ . Given  $\Delta H = H_0 - H_M$  (10), an effective strategy for social change should be increasing information entropy  $H_M$  in the pro-change power's network patterns (9) and reducing  $\Delta H$  (i.e. amount of memory; quantitative theory of adaptive animal behavior [1-3]). Put another way, this model suggests that activities for complex systems social change should focus on informational and ideological confrontations via peaceful means, rather than forceful struggle. In order to test the validity of this model, the question of what to use to ascertain the multiple possible social change outputs is extremely important. Within developed nations, the foundation of society is a representative democracy composed of a party government. By way of addressing one social (change) output, we selected the share of votes received by political parties; their rates of change were then aggregated and analyzed over a number of years. The rate of change in the vote share is the absolute difference in vote share between one election and the next at fixed intervals. In terms of the model, by examining not the size of the rate of change, but rather the distribution's exponents, the agonistic structure (real power relationship) can be ascertained. In this paper we focus on the national elections held in Japan after World War II. In Japan many political parties were repeatedly launched, merged and divided during a period of seventy years after the war and at present eleven parties exist with seats in the Diet of Japan which consists of Upper and Lower Houses. Among the parties only the Japanese Communist Party, which is the third largest opposition party, and the Liberal Democratic Party, which is now in power, have been maintaining the unity over half a century. The other parties including the present largest opposition party, the Democratic party of Japan, have been founded recently. Hence, one understands that it is appropriate to choose the above two long-lived parties in the analysis of the agonistic structure in Japan. Specifically, we used national (1947-1980) and proportional-representative (1980-2010) Upper House elections, which are held regularly every three years and in which all of Japan is treated as a single constituency. First we will examine the Japanese Communist Party (JCP), which has long operated under the same name and has the largest sample size, and then the Liberal Democratic Party of Japan (LDP). Finally, we will look at the relationships between them.

## RESULTS

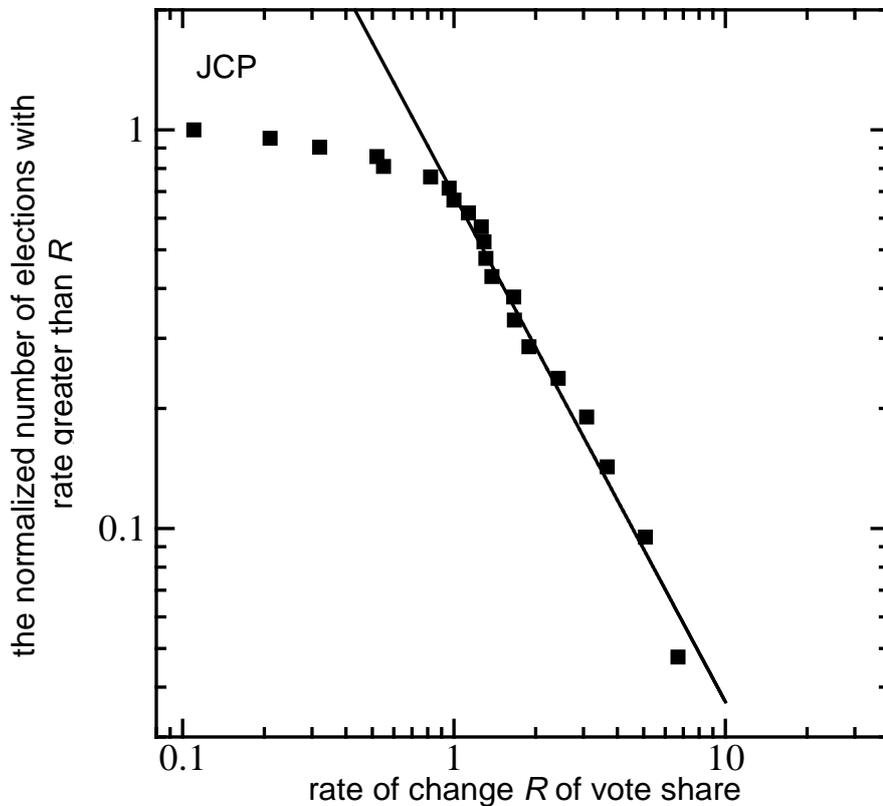
### JCP VOTE SHARE

Figure 1 shows the trend of the JCP's vote share in Upper House elections from 1947 to 2010. Twenty-two elections have been held during this time. The vote share changes dramatically, reflecting the complex political, economic, and cultural conditions of the times. At its peak in 1998, the JCP "surged" when the Hashimoto government suffered a major defeat after raising the consumption tax to 5 %, causing economic deterioration. As a dynamic output of social change, analysis is performed with respect to the rate of change  $R$  of the vote share<sup>2</sup>.

Figure 2 shows a double logarithmic plot with the rate of change  $R$  on the horizontal axis, and the normalized number of elections with a rate of change greater than  $R$  on the vertical axis. Surprisingly, excluding five points of data in which the rate of change is small, the points fall along a single line and clearly exhibit a specific power-law distribution with an exponent of 1.27. As there seems to be nothing special about the times when the excluded five data points occurred, they can be regarded as noise from the view point of social change (since we are focusing on large  $R$  values). This result shows that there is a clear law governing fluctuations in vote share determined by complex, diverse factors in each era; in the case of the JCP's vote share, there exists a specific social change equation (power law distribution with the exponent



**Figure 1.** The trend of the JCP's vote share in Upper House elections from 1947 to 2010.



**Figure 2.** A double logarithmic plot with the rate of change  $R$  of vote share on the horizontal axis, and the normalized number of elections with a rate of change greater than  $R$  on the vertical axis. The power exponent  $D_M$  obtained in terms of the least squares method is 1,27.

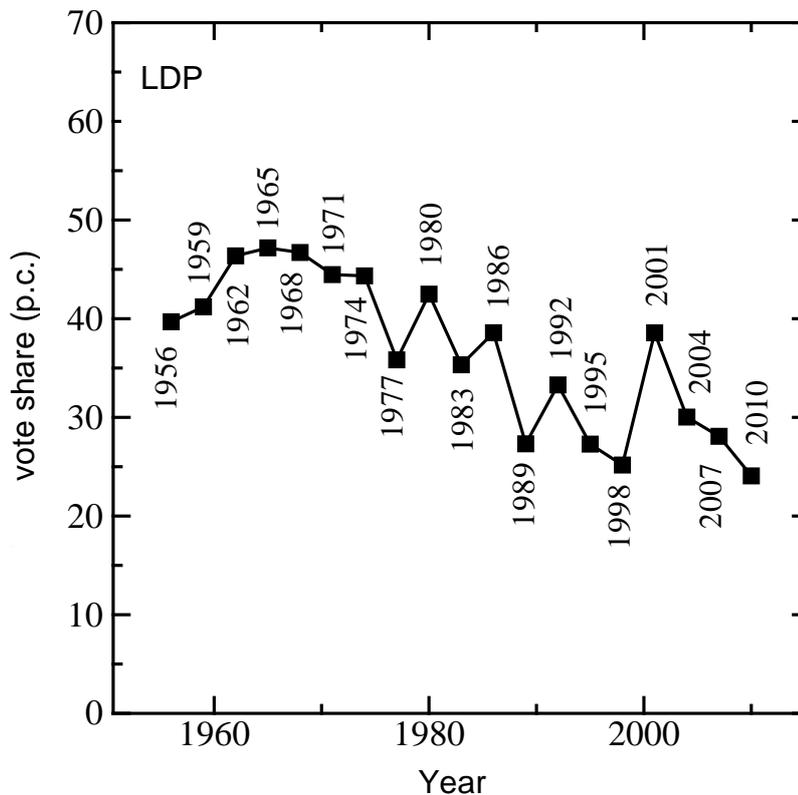
1,27). Large fluctuations (e.g. 1998) conform to this equation and are not anomalous surges. That is to say, the election results, including the surges, have not changed the real power relationship as expressed by the power exponent 1,27.

### LDP VOTE SHARE

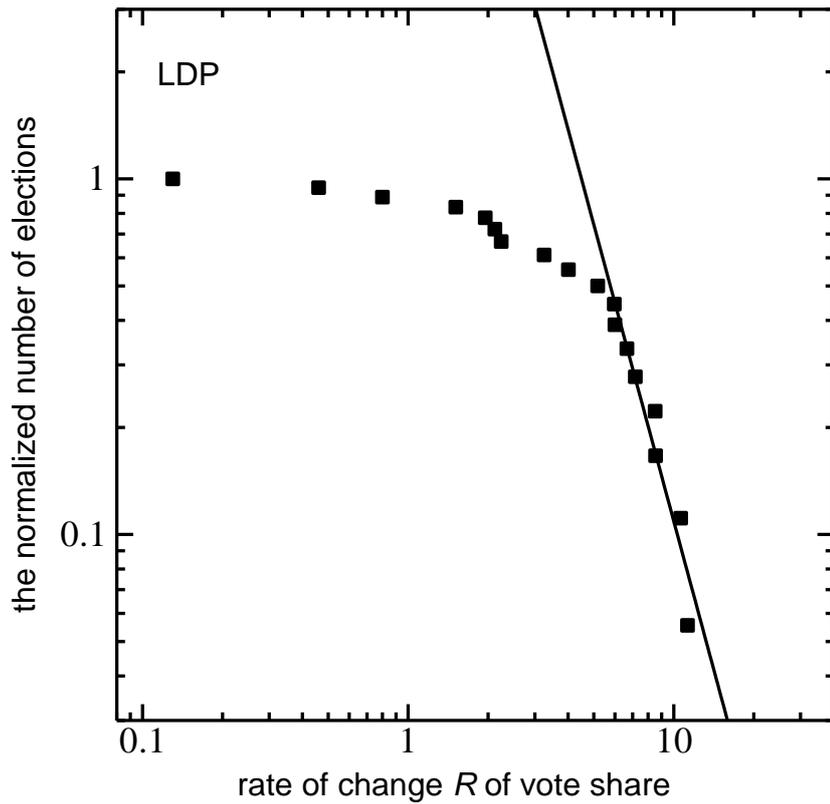
Since its formation in 1955 through a conservative alliance, the LDP has lost twice to opposition parties, falling out of power for brief periods. Otherwise, it has occupied a central position as a consistently pro-establishment power. Figure 3 shows the trend of the LDP's vote share in 19 elections from 1956, the year after its formation, to 2010. Figure 4 shows a double logarithmic plot with the rate of change  $R$  on the horizontal axis, and the number of elections with a rate of change greater than  $R$  on the vertical axis. The majority of data points, 10 of 18, diverge from a straight line. Here, the power exponent is 2,27. Next, a single logarithmic plot of the same data is shown in Figure 5. All data falls along a straight line, and displays exponential distribution. This is in contrast with the JCP's power-law distribution.

### RELATIONSHIP OF PARTIES' VOTE SHARE

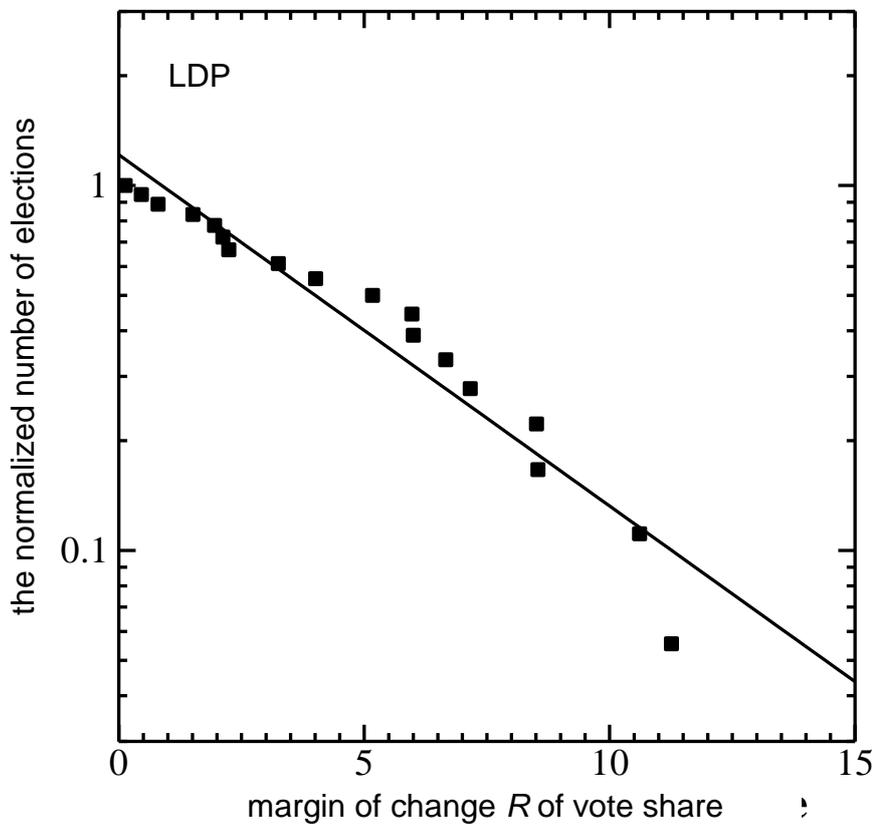
Figure 6 plots the vote share of the JCP (horizontal axis) and LDP (vertical axis), showing the correlation between the two. It indicates a clear negative correlation. When the JCP vote share rises, the LDP share falls; the reverse is also true. The isolated point to the bottom-right represents data from the 1998 Hashimoto cabinet Upper House election, but aside from this point, the correlation coefficient does not fluctuate. This supports the validity of the assumption of the complex systems social change model corresponding to Eq. (4) in Appendix.



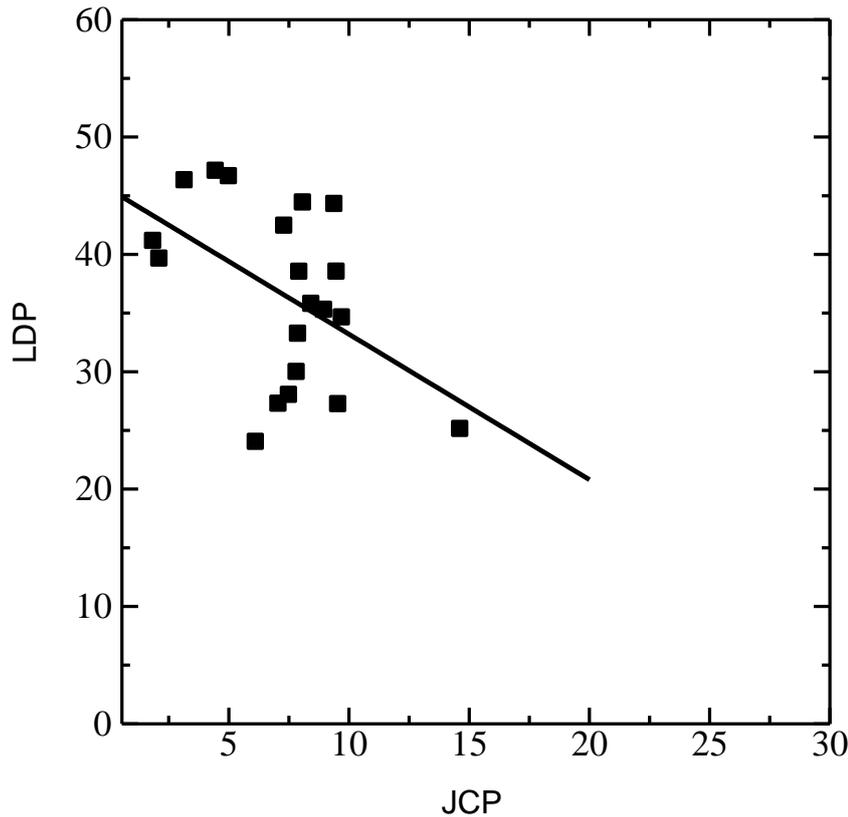
**Figure 3.** The trend of the LDP's vote share in 19 elections from 1956, the year after its formation, to 2010.



**Figure 4.** A double logarithmic plot with the rate of change  $R$  on the horizontal axis, and the number of elections with a margin of change greater than  $R$  on the vertical axis. The power exponent in terms of the least squares method is 2,27.



**Figure 5.** A single logarithmic plot of the same data as shown in Figure 4.



**Figure 6.** Plots of the vote share of the JCP (horizontal axis) and LDP (vertical axis). The coefficient of the correlation value is  $-0,48$ .

## DISCUSSION

The JCP displays a definite power law and large changes in its vote share; within the model, it exhibits the characteristics of a pro-change power, which corresponds with its actual role. In contrast to the JCP, the LDP, which is regarded as the establishment party, exhibits an exponential distribution, and has a relatively stable rate of change with minor fluctuations. Viewed theoretically from the establishment side, when  $\beta$  is small, and the pro-change power's resistance probability  $Q(\beta N)$  is small, this leads to the stable continuation of the exponential distribution (8). However, as indicated by the fallout from the Lehman collapse and the severe deflationary economic downturn in Japan, the buildup of discrepancies in the basis of Japanese society is reaching an extreme level. In order to break free from this stagnant era and open up future prospects, new strategy for social change is needed. Currently, the country is approaching a dramatic transition from the state of social change that has existed up until now (exponent  $D_M = 1,27$ ) to a state of real criticality, the most probable for social change ( $D_M = 1$ ). The model suggests that a decrease of the power exponent  $D_M$  (1,27) changes the agonistic structure greatly, leading to an increase in the change output  $R$ .

In order to decrease the power exponent  $D_M = \alpha_M/\beta$  (slope),  $\alpha_M \sim \alpha_0 \exp(\Delta H)$ , thus a decrease in  $\Delta H = H_0 - H_M$  will be effective. Consequently, increasing the information entropy  $H_M$  (9) of the pro-change power's network pattern should be effective for real social change. In other words, the model suggests an "information entropy enhancement" strategy. This raises the question of what it means to increase information entropy in the pro-change power's network pattern. First it means the expansion of the uncertainty and variety of information. It denies the concentration of specialized information, and expands a united front network composed of diverse singular issues (single points) – i.e. letting "a hundred schools of thought contend".

Moreover, in the quantitative theory of adaptive animal behavior,  $\Delta H$  is the amount of memory, so reducing it should facilitate new information creation, sharing and dissemination, and departure from experientialism, the revelation of hidden social realities, etc. Stirrings such as the Arab Spring, Occupy Wall Street, and protests at the Prime Minister's residence (of Japan) are real examples of a new social change movement abroad and in Japan. Therefore, the central issues for social change movements are the informational and ideological struggle and the construction of a diverse social change network.

On the other hand, approval ratings for governments showed exponential distributions [4]. The results may indicate from our model that the approval groups examined were not in the 'critical state' of mutual antagonism and conflict between approval and disapproval groups.

The above is a quantitative discussion of social change based on models. It does not address the quality (contents) of change, such as what will change or what kind of society will be created. The united front strategy based on democratic reform theory [5] essentially and qualitatively defines the object of change above all else. The quantitative "information entropy enhancement" strategy proposed in this research is not inconsistent with the united front strategy; the two strategies can compatibly supplement one another. A deep analysis of the relationship between the two will be necessary in the future.

## **POSTSCRIPT**

During the Upper House election on July 21, 2013, the JCP's vote share surged from 6,1 % for the previous election to 9,7 %. This can be seen as the result of an, albeit unintentional, information entropy enhancement strategy of a diverse united front of singular issues (consumption tax, constitutional reform, nuclear power, the Great East Japan Earthquake, Trans-Pacific Partnership, American military bases, social security, etc.). However, upon analysis via the model for complex systems social change proposed by this research, there was almost no change in the power exponent, so this was not enough to alter the real power relationship of the agonistic structure. This also indicated that the JCP's medium range vote share goal of 20 % should achieve a significant change in the power relationship.

## **CONCLUSIONS**

We constructed a complex systems 'criticality' model, applied it to social change, and examined its reliability and validity. Social change manifests in the 'critical state' of mutual antagonism and conflict between the pro-change and pro-establishment powers. The model derived a power law distribution of the output of social change, the characteristics of a pro-change power, while a pro-establishment power displayed an exponential distribution. The power exponent reflects the agonistic structure (real power relationship) of both powers. The validity of the model was verified by examining vote shares of parties in Japan. Based on the examination, we propose a new quantitative strategy "information entropy enhancement" for effective change of the exponent, and therefore, for social change.

## **REMARKS**

<sup>1</sup>This hypothesis is analogous to the Boltzmann distribution, but it is also a solution of  $dI(N)/dN = \alpha I(N)$  which is often used in various growth phenomena. Recently, it was shown that the model holds when power law distribution  $N^\alpha$  is hypothesized. Because  $N^\alpha = \exp(\alpha \ln N) = \exp(\alpha M)$  when  $\ln N = M$ , all procedures in the model are the same for  $M$ . This result suggests that the assumption of exponential function in equation (1) is not essential; a model can potentially be realized if there exists an increasing function.

<sup>2</sup>The rate of change  $R$  in the vote share is the absolute difference in vote share between one election and the next.

## APPENDIX: COMPLEX SYSTEMS ‘CRITICALITY’ MODEL

### GENERAL THEORY OF THE ‘CRITICALITY’ MODEL IN COMPLEX SYSTEMS

Criticality refers to “a state in which two powers both act towards creating a different state, and in so doing both achieve a state of precise balance” (Butsurigaku Jiten in Japanese). However, research [6] indicates that in complex systems multiple metastable self-organized critical states are formed. The word ‘criticality’ is used here in quotes to distinguish it from the precisely balanced criticality of simple systems physics. Hereafter the term will be used as-is, without quotes, to indicate criticality in complex systems, unless otherwise noted.

The criticality model of complex systems is predicated on a fundamental idea of complex systems, namely that “complex system dynamics and its output (quantitative, qualitative) are defined by a network pattern of diverse mutual interactions (physical, informational, social) working among  $N$  active elements that are related to output, within the elements that comprise the system”. This view of complex system dynamics has come to be shared in recent years. However, this model is limited to the quantitative aspects of the system, and is based on the assumption that the output is defined by the existence probability  $P(N)$  of the number of network patterns  $I(N)$ . In cases where a positive tendency and opposing anti-tendency are observed on that network, the existence probability  $P(N)$  of the number of network patterns  $I(N)$  expressing the positive tendency and the existence probability  $Q(N)$  that its anti-tendency can take, describe the two conflicting tendencies in terms of probability theory. These do not touch upon the details of mutual interactions between elements. As shown in the next section, the output  $R$  of a critical state in a system with two opposing, mutually-antagonistic tendencies, logically leads to a power-law distribution. As far as we know, there are no previous examples of power-law derivation in complex systems based on the above assumption, and thus this assumption's validity must be tested by applying it to various complex phenomena. When doing so, it must be acknowledged that an output system will form in which qualitatively different outputs will exist even for a single complex system and each of these could in turn manifest multiple quantitative outputs. For example, in animal behavior, these could be the duration of the behavior [1-3], its intensity, scale, etc.; in earthquakes, the earthquake's energy or magnitude [7], duration, fault surface, etc.; in a conflict, the number of casualties [8], length of the conflict, etc.; and in social change, these could be the percentage of votes taken by the pro-change power in the present paper, participants in protests, etc. The question of what to capture as the output is extremely important when examining the model, but only the quantitative aspects of the output can be analyzed/examined.

### DERIVATION OF THE CRITICALITY MODEL OF COMPLEX SYSTEMS

Suppose  $N$  is the number of active constituent elements that determine a complex system's dynamics. Assume a number of network patterns  $I(N)$  of mutual interactions between those elements in the following equation (and see remark 1)

$$I(N) \propto e^{\alpha N} . \quad (1)$$

$\alpha$  is a parameter that expresses rate of growth with respect to  $N$  of the number of patterns  $I(N)$ . The existence probability  $P(\alpha N)$  of the number of patterns  $I(N)$  is expressed as,

$$P(\alpha N) = \frac{e^{-\alpha N}}{\sum_{N=0}^{N_C} e^{-\alpha N}} , \quad (2)$$

and so the largest number of active elements is  $N_C$ . In the system's critical state, the existence probability  $Q(\beta N)$  of the number of patterns of anti-tendencies, acting in opposition and

antagonistically to this network of positive tendencies, is expressed as a similar exponential function, with the parameter  $\beta$  generally differing from  $\alpha$

$$Q(\beta N) = \frac{e^{-\beta N}}{\sum_{N=0}^{N_c} e^{-\beta N}} \quad (3)$$

$P(\alpha N)$  and  $Q(\beta N)$  are common simple function types that have interesting and somewhat complex characteristics. Both are monotonically decreasing functions of  $N$ , but if one examines the dependency on  $\alpha$  or  $\beta$ , when  $\alpha$  is sufficiently small,  $P(\alpha N)$  is also small ( $\sim(N_c + 1)^{-1}$ ), but largely depends on  $N$ . That is, when  $0 < N < N_c/2$ , a single extreme value exists for a given  $\alpha$ , and when  $N_c/2 < N$ , it monotonically decreases with respect to  $\alpha$  (unpublished).

Next, these two probabilities express two opposing tendencies, and define the output system  $R$  that is produced from the conflicting critical state; this results from the following. Initially, as the output  $R$  will probably decrease as anti-tendencies strengthen, it is assumed to be inversely proportionate to the probability  $Q(\beta N)$ , so  $R$  is defined by  $\beta N$ .

$$\frac{1}{R} = \frac{1}{R_0} Q(\beta N) = \frac{1}{R_0} \frac{e^{-\beta N}}{\sum_{N=0}^{N_c} e^{-\beta N}} \quad (4)$$

Moreover, if  $R$  is sufficiently large and integral (and continuous) representation is employed,  $\phi(R)$  is the probability density function of  $R$ . Through variable transformation by  $R$  of the existence probability  $P(\alpha N)dN$ , the probability  $\phi(R)dR$ , which expresses the structure of the output system  $R$ , is acquired (probability conservation). That is, given  $\phi(R)dR = P(\alpha N(R))(dN/dR)dR$ ,

$$\frac{dN}{dR} \propto \frac{1}{\beta} e^{-\beta N}, \quad (5)$$

which follows from Eq. (4), then

$$\phi(R) = P(\alpha N) \frac{dN}{dR} \propto (e^{-\beta N})^{\alpha/\beta} (1/\beta) e^{-\beta N} \propto (R/R_0)^{(-\alpha/\beta-1)} \quad (6)$$

Integrating this (cumulative distribution)

$$\Phi(R) = \int_R^{\infty} \phi(R)dR \propto (R/R_0)^{-\alpha/\beta} \quad (7)$$

yields a power-law distribution expressing the structure of the complex output system  $R$ . With regards to  $\alpha$  and  $\beta$ , opposite tendencies are expressed by the numerator and denominator in  $D_M = \alpha/\beta$ . However, if  $\beta$  is sufficiently small, this will become an approximately exponential distribution:

$$\Phi(R) \propto \exp(-\alpha R / R_0 \beta) \quad (8)$$

The complex system is simultaneously a complex material system and a complex informational system. In order to quantitatively handle the complex system, we will introduce the complex system network's information entropy (mean information amount)  $H_M$  to the model. Shannon's definition of information entropy  $H_M$  is expressed as a discrete equation, but it can be approximately expressed as the following integral representation.

$$\int_0^{N_c} dN [-\log P(\alpha_M N)] P(\alpha_M N) \equiv H_M \quad (9)$$

Here,  $\alpha_M$  and  $P(\alpha_M N)$  are the parameter and existence probability, incorporating the information effect. If the change in information entropy of the complex system network  $\Delta H$  is

$$\Delta H = H_O - H_M \quad (10)$$

( $H_O$  is the maximum value of  $H_M$  at the time  $\alpha_M$  changes), then when  $N_C$  is large [3]:

$$H_S \approx 1 - \log \alpha_S (s = 0 \text{ or } M). \quad (11)$$

$H_S$  is expressed by  $\alpha_S$ , therefore:

$$\alpha_M = \alpha_O e^{\Delta H}. \quad (12)$$

Consequently

$$D_M = \frac{\alpha_M}{\beta} \quad (13)$$

Thus,  $D_M$  is shown to change depending on  $\Delta H$ . However, it may be necessary to incorporate the information effect in the probability  $Q(\beta N)$  for specific complex phenomena.

## SIGNIFICANCE OF MODEL

This model illustrates one general expression of the laws of transformation and criticality, which comprise the foundation of the developmental/evolutionary processes of complex systems. Moreover, it points to the importance of the meaning of and changes in the power exponent  $D_M$  in complex system dynamics that express the power-law, suggesting that in these changes, the contribution of  $\Delta H$  is (exponentially) large. In fact, in the adaptive animal behavior model [1-3], which is the basis for this model, it has been clearly demonstrated, through five independent experiments, that changes of the power exponent  $D_M$  depend on amount of memory ( $\Delta H$ ) and gustatory stimulus ( $\beta$ ).

Up until now, it seems that many complex systems models are used to provide feasible, a posteriori explanations, or to tout new and different interpretations [9]. Perhaps we are merely ignorant of their existence, but there seem to be almost no widely effective and potentially predictive theories (models). In addition, there have been almost no examples of research that directly explore changes in the exponents of power law within complex systems science. Rather, the research attempts to evaluate the universality of power exponents beyond differences of object. In elucidating the significance of power exponent dynamics within the same object, this model may clarify system structural dynamics, which are at the root of output fluctuations in complex systems, and suggests the potential of structural “change” and “control”.

## ACKNOWLEDGEMENTS

We thank Drs. Hiroshi Hamada, Satoshi Ihara, Hitoshi Kono, Yatsuhisa Nagano, Hiroshi Onishi and Reiji Sugano for useful discussions.

## REFERENCES

- [1] Shimada, I. and Koganezawa, M.: *Animal Behavior as a Complex Adaptive System*. Comparative Physiology and Biochemistry **28**, Suppl. (Abs.) 2011,
- [2] Shimada, I.; Kawazove, Y. and Hara, H.: *A Temporal Model of Animal Behavior Based on Fractality in the Feeding of Drosophila Melanogaster*. Biological Cybernetics **68**(6), 477-481, 1993, <http://dx.doi.org/10.1007/BF00200806>,
- [3] Koganezawa, M.; Hara, H.; Hayakawa, Y. and Shimada, I.: *Memory Effects on Scale-free Dynamics in Foraging Drosophila*. Journal of Theoretical Biology **260**(3), 353-358, 2009, <http://dx.doi.org/10.1016/j.jtbi.2009.06.018>,

- [4] Honjo, H.; Sano, M.; Miki, H. and Sakaguchi, H.: *Statistical Properties of Approval Ratings for Governments*.  
*Physica A* **428**, 266-272, 2015,  
<http://dx.doi.org/10.1016/j.physa.2015.02.047>,
- [5] Japanese Communist Party: *Program of the Japanese Communist Party*.  
<http://www.jcp.or.jp/english/23rd-congress/program.html>,
- [6] Bak, P.: *How Nature Works: The Science of Self-organized Criticality*.  
Springer, New York, 1999,
- [7] Gutenberg, B.C. and Richter, C.F.: *Frequency of Earthquakes in California*.  
*Bulletin of Seismological Society of America* **34**, 185-188, 1944
- [8] Richardson, L. F.: *Variation of the Frequency of Fatal Quarrels with Magnitude*.  
*Journal of American Statistical Association* **43**(244), 523-546, 1948,  
<http://dx.doi.org/10.1080/01621459.1948.10483278>,
- [9] Watts, D.J.: *Everything is Obvious: Once You Know the Answer*.  
Crown Business, New York, 2011.

# THE CONCEPT OF FREE WILL AS AN INFINITE METATHEORETIC RECURSION

Hashim Hanaan<sup>1</sup> and Srikanth Radhakrishna<sup>2, \*</sup>

<sup>1</sup>Sibylgaze, Industry Intelligence GmbH  
Bangalore, India

<sup>2</sup>Poornaprajna Institute of Scientific Research  
Sadashivnagar, Bangalore, India

DOI: 10.7906/indecs.13.3.2  
Regular article

*Received:* 2 April 2015.  
*Accepted:* 10 June 2015.

## ABSTRACT

It is argued that the concept of free will, like the concept of truth in formal languages, requires a separation between an object level and a meta-level for being consistently defined. The Jamesian two-stage model, which deconstructs free will into the causally open “free” stage with its closure in the “will” stage, is implicitly a move in this direction. However, to avoid the dilemma of determinism, free will additionally requires an infinite regress of causal meta-stages, making free choice a hypertask. We use this model to define free will of the rationalist-compatibilist type. This is shown to provide a natural three-way distinction between quantum indeterminism, freedom and free will, applicable respectively to artificial intelligence (AI), animal agents and human agents. We propose that the causal hierarchy in our model corresponds to a hierarchy of Turing uncomputability. Possible neurobiological and behavioral tests to demonstrate free will experimentally are suggested. Ramifications of the model for physics, evolutionary biology, neuroscience, neuropathological medicine and moral philosophy are briefly outlined.

## KEY WORDS

free will, uncomputability, infinite recursion, Jamesian two-stage model, quantum indeterminism

## CLASSIFICATION

JEL: I31

PACS: 89.90.+n

\*Corresponding author, *η*: srik@poornaprajna.org; +91 080 23611836;  
Poornaprajna Institute of Scientific Research, 4 Sadashiv nagar,  
Bangalore – 560 080, Karnataka, India

## INTRODUCTION

Free will (FW) is a concept in philosophy that refers to the putative capacity of a human agent to control her behavior by choices made by an act of will, on basis of her personal motives, convictions and intentions. The concept rests on the belief that human behavior is not fully determined by external causes. Moreover, her motives, convictions and intentions are themselves not determined by fully external causes, but self-determined. From a common sense perspective, we feel we are free in making decisions. Yet it continues to be debated, even after centuries of argumentation, how to coherently define free will and whether it exists in Nature [1].

FW can be regarded as freedom from some constraint: exactly what that constraint is, has remained moot. There are two broad philosophical positions on FW: *Incompatibilism*, which holds that the relevant constraint is determinism, and *Compatibilism*, which holds that determinism is irrelevant to the definition of free will, and that determinism and FW are compatible.

Two divergent Incompatibilist positions are *Hard Determinism*, which regards FW as false and determinism as true, and (metaphysical) *Libertarianism*, which regards determinism as false and FW as true. A FW *Skeptic* is an Incompatibilist who goes farther than a Hard Determinist and denies that FW is even a coherent concept. From the Skeptic perspective, indeterminism no more allows an agent control and self-determination over actions than does determinism. Since determinism and indeterminism are the only two logically possible causal primitives, this view holds that very concept of FW is meaningless. This Skeptic stand is sometimes called the dilemma of determinism or the standard modern argument against FW.

The libertarian may imagine that housed in her brain is an immaterial agency (such as a soul or homunculus) that somehow transcends the cause and effect law that holds elsewhere in Nature. This is a logically tenable defense of Libertarianism against Hard Determinism, even if an adherent of the latter would deny the existence of such immaterial agencies. But even the soul offers Libertarianism no protection against the Skeptic, because if the soul's choices transcend causality, they must be random, or results of random properties, again making the case for FW bleak!

The present work hopes to convince the FW Skeptic that a coherent metaphysical account of FW is possible. Of importance to our work is using an idea similar to that used by Tarski to define the concept of truth, except that here it is applied to causality. We think that FW is a causal primitive different from both determinism and indeterminism in being *metatheoretic*. Our attempt to define it shows it to be a form of causation that straddles endless levels of causality, *provided* we wish to form a coherent scientific narrative that fits in with our intuitive sense of self, freedom and responsibility.

This article is structured as follows. A version of the Jamesian two-stage model is proposed in Section 2, which would be useful for our further discussion. Although such a model has been considered as the defense of FW against the dilemma of determinism, we point out in Section 3 that the model fails, because the dilemma can be recursively resurrected. In Section 4, we show that an infinite recursion of Jamesian two-stage models restores a measure of protection for libertarianism against the dilemma. This model is adapted to define a Rationalist-Compatibilist FW. The connection of FW to uncomputability [2] and Tarskian undefinability for formal truth [3] are outlined in Section 5. We then discuss the consequences of the model for neurobiology in Section 6, before concluding in Section 7.

## THE TWO-STAGE MODEL, 2S: A NEW HOPE

The two-stage model [4], introduced in its original form by W. James [5], and proposed in various forms by a number of researchers, is intended to defeat the dilemma of determinism.

The model posits that FW is a two-stage process: first there is freedom at the lower stage, and then there is will at the higher stage, which makes a choice. We present further in the text a version of it, which we call ‘2S’, with several changes in the details, as described in the following two subsections.

### PHYSICAL STAGE L<sub>0</sub> (FREEDOM)

At the moment an agent’s attention is drawn to a conflict situation, during a short time span, called the *selection window*, in a localized region of brain, which we refer to as the *free-will oracle*, probably in the pre-frontal cortex, the physical laws  $W_0$  are put on a “causally open” mode. In preparation of entry into this mode, alternative options  $x_0$ , described by probability distribution  $P_0$ , are generated in the agent’s brain. The physical laws  $W_0$  only determine  $P_0$ , but do not entirely fix the eventual choice  $x_0$ . And the selection is not completely determined by the past history of the *physical* universe. At the close of the selection window,  $W_0$  in FW oracle is re-set to the “causally closed” mode, and the choice  $x_0$  that is available on the FW oracle’s “register” at that moment is expressed as the agent’s action.

If the agent makes the choice mechanically, without a focused exertion of her will, then there is no mental causation influencing the choice (as described below), and at the close of the selection window, a random  $x_0$  is selected according to probability distribution  $P_0$ .

### METAPHYSICAL STAGE L<sub>1</sub> (WILL)

If the agent decides to exert her will, then her choice must not be random but instead reflect the desires, intentions and beliefs characteristic of the agent. These properties constitute her *cognitive private space*, and may not have a full representation in the physical level  $L_0$ . Thus the will produces generally a *deviation* from  $P_0$ .

A simple way to describe this situation mathematically is by:

$$x_0 = \lambda_1(P_0), \tag{1}$$

where  $\lambda_1$  is the “will function” that encodes the properties of her cognitive private space. Function  $\lambda_1$  takes note of  $P_0$ , but is *not* part of  $W_0$ , because if it were, then there would be causal closure in the physical, and hence no freedom at  $L_0$ . Instead,  $\lambda_1$  is determined by the laws  $W_1$ , which extend  $W_0$  to  $L_1$ .

The fact that the causal openness at  $W_0$  is replaced by causal closure of  $W_1$  in the consolidated system in  $L_0 + L_1$  suggests that these two levels can be ordered in a *causal hierarchy*, which we represent by the expression:  $L_1 < L_0$ . The ordering expresses that  $L_1$  causally precedes  $L_0$ .

### 2S AND THE DILEMMA OF DETERMINISM

In the absence of action by the will, the FW oracle can be considered as a *probabilistic* input/output machine. To underscore this, sometimes we will refer to such indeterministic behavior as *freedom-without-will*. We introduce a level  $L_{0.5}$  as essentially  $L_0$  equipped with a source of pure randomness. Instances of freedom-without-will have their causal closure in  $L_0 + L_{0.5}$ , but in this case there is no deviation from  $P_0$ . When the will is exerted, the random variable  $X_0$  representing free choice deviates from  $P_0$ , and free will transcends physical causality.

The attempted defense of FW against the dilemma of determinism using model 2S would be: FW is not deterministic because of freedom at the physical stage. Nor is it random because the eventual choice is self-determined, being fixed by personal preferences via  $\lambda_1$ .

Though not the definitive word on FW, the model 2S brings the new insight that FW is a new *causal primitive*, different from determinism and indeterminism. While the latter two can be associated with a causally closed lawfulness, and thus defined on a single causal level (say  $L_0$ ), FW cannot. More of this further in the text.

## FREE WILL IN 2S AND FORMAL TRUTH

As we indicated in Ref. [6], the 2S feature of introducing the metaphysical level to define FW on the physical level parallels Tarski's use [3] of a metalanguage in order to consistently define the concept of arithmetic truth in an object language. Tarski showed that without careful separation of the two levels, one would end up with logical antinomies like the liar's paradox "This statement is false" (which is true iff it is false).

Similarly, if we fail to separate the "free" and "will" stages in free choice, we will err in reducing the choice to determinism or randomness. We believe that part of the difficulty in understanding the nature of FW is due to a lack of appreciation for this level separation.

With this in mind, we may consider  $L_0$  as the object stage, and  $L_1$  as the metastage. The free-willed agent with 2S structure will be designated  $F_1$ . A deterministic physical system, which is causally closed at level  $L_0$  is designated  $F_0$ , while an indeterministic quantum physical system, which has  $L_0$  freedom-without-will will be denoted  $F_{0.5}$ .

## DILEMMA OF DETERMINISM REVISITED

Although the model 2S resolves the dilemma of determinism after a fashion, still the dilemma can be resurrected at level  $L_1$ . To see this, note that the two-stage agent  $F_1$ , taken as a whole, is deterministic, in view of Eq. (1), and must thus lack FW. This is just Schopenhauer's argument, who picturesquely said, "Man can do what he wants but he cannot will what he wants" [7], in his prize-winning essay in response to the challenge posed by the Royal Norwegian Academy of Sciences in 1839<sup>1</sup>.

To reinstate free will, we allow the will function  $\lambda_1$  to be freely chosen, i.e., we apply 2S to the selection of  $\lambda_1$ . We extend the FW oracle from the physical level  $L_0$  to the metaphysical  $L_1$ . Within the selection window, the laws  $W_1$  governing level  $L_0 + L_1$  become momentarily causally open in the region of the FW oracle. We thus have *freedom of the will* at  $L_1$  whereby the  $L_0 + L_1$  (i.e., physical-metaphysical) history of the universe does not fix  $\lambda_1$ .

We introduce a metalevel  $L_2$ , where a higher-order will function  $\lambda_2$  ("will free will"), provides causal closure, by deterministically selecting a particular  $\lambda_1$ . We can think of  $\lambda_2$  as representing a deeper aspect of the agent's character or disposition and as determining the type of the will function  $\lambda_1$  that she will select. Function  $\lambda_2$  comes from a further interior layer of the cognitive private space of the agent. By adding an element of spontaneity and self-determination, it thwarts the selection of  $\lambda_1$  from being modelled as an input/output process at level  $L_1$ . In the metatheoretic representation, this is tantamount to treating  $F_1$  as the object system, and an  $L_2$ -aspect as the metasystem. A simple mathematical way to represent this is by:

$$\lambda_1 = \lambda_2(P_1), \quad (2)$$

where  $P_1$  is a *probability function* encoding  $L_1$ -preferences according to laws  $W_1$ . If the higher-order will  $\lambda_2$  is not exerted, then  $P_1$  would describe the indeterministic selection of will  $\lambda_1$ . But if this "will free will" is exerted, then there will be deviations from  $P_1$  in the selection of  $\lambda_1$ . Substituting Eq. (2) into Eq. (1), we obtain:

$$x_0 = \lambda_1(P_0) = [\lambda_2(P_1)](P_0) \equiv \lambda_2^*(P_1, P_0). \quad (3)$$

The second-order will  $\lambda_2$  selects a first-order will  $\lambda_1$  depending on  $P_1$ , and then  $\lambda_1$  selects  $x_0$ . The extended causal hierarchy is  $L_2 < L_1 < L_0$ , whereby  $L_2$  causally precedes  $L_1$ , which in turn precedes  $L_0$ .

Here precedence refers not just to chronology but to recursion depth in the agent's cognitive private space. It does not refer to an *earlier* event on the same causal level, but to *deeper* or higher-order cause. For example, if  $\lambda_1$  inclines Alice to give Bob a gift, then  $\lambda_2$  could be her character trait of wanting to help him. A yet deeper cause ( $\lambda_3$ ) would be wanting for good to him or people she likes. And so on.

We denote by  $F_2$  the type of agent for who there is causal closure at  $W_2$ , and by  $F_{1,5}$  one for who the will function  $\lambda_1$  is selected indeterministically. Although the existence of  $F_2$  resolves the dilemma of determinism on  $L_1$ , still the dilemma can be resurrected at level  $L_2$ . This is evident seeing that the action of  $\lambda_2^*$  in Eq. (3) is deterministic.

To reinstate free will at  $L_2$ , we recursively apply  $2S$ . We make  $L_2$ -law  $W_2$  at the FW oracle also causally open within the selection window. Let the freedom of  $\lambda_2$  be described by probability function  $P_2$ , determined by  $W_2$ . Causal closure occurs at  $W_3$  via higher-order FW  $\lambda_3$  from an even deeper aspect ( $L_3$ -aspect) of the agent's cognitive private space. We have

$$\lambda_2 = \lambda_3(P_2), \tag{4}$$

Substituting Eq. (4) into Eq. (2), we obtain:

$$x_0 = [\lambda_2(P_1)](P_0) = [[\lambda_3(P_2)](P_1)](P_0) \equiv \lambda_3^*(P_2, P_1, P_0). \tag{5}$$

We extend the causal hierarchy as  $L_3 < L_2 < L_1 < L_0$ . An agent with causal closure of her choosing process in  $W_3$  is denoted  $F_3$  and one with freedom-without-will on  $L_2$  (i.e.,  $L_2$ .indeterministic) is denoted  $F_{2,5}$ .

Although the existence of  $F_3$  appears to resolve the dilemma of determinism on stage  $L_2$ , still the dilemma can be resurrected at stage  $L_3$  since  $F_3$ , in view of Eq. (5), can be considered as a deterministic system. This may be seen as a further higher-order extension of Schopenhauer's argument.

To prevent the dilemma of determinism at stage  $L_3$ , we introduce a  $2S$  model on top of this stage, with  $F_3$  as the object system, and  $L_4$  as the metastage. But then the resultant  $F_4$  will be deterministic. We require  $L_5$  to obtain a  $2S$  model for  $F_4$  to avoid the dilemma of determinism at level  $L_4$ . Continuing this trend indefinitely, at level  $L_n$ , where  $n$  is any positive integer, we have:

$$\lambda_{n-1} = \lambda_n(P_{n-1}). \tag{6}$$

Substituting this recursively into lower levels into Eq. (3), we obtain:

$$x_0 = [[\cdots [[\lambda_n(P_{n-1})](P_{n-2})] \cdots ](P_1)](P_0) \equiv \lambda_n^*(P_{n-1}, \cdots, P_0). \tag{7}$$

Evidently, we can still resurrect the dilemma of determinism at level  $L_{n+1}$  no matter how large  $n$  is, since the fact of  $\lambda_n^*$  in Eq. (7) is a deterministic function. Thus, the problem posed by this dilemma does not disappear for agent  $F_n$  but is merely postponed.

In response to this seemingly insurmountable difficulty posed by the recursive version of the dilemma of determinism, the only option to save libertarianism seems to be to let a free-willed agent be an infinite-stage entity,  $F_\infty$ . What this means is that, for any *finite* integer  $n$ , the laws ( $W_n$ ) of cause and effect at the stage  $L_n$  will lack causal closure in the region of the extended FW oracle during the selection window, and the selection of will  $\lambda_n$  cannot be modelled as a probabilistic or deterministic input/output system at stage. The causal closure for  $W_n$  will come through a higher-order cause  $\lambda_{n+1}$ , which is determined by the  $L_{n+1}$ -aspect of the cognitive private space. Therefore, the agent's choice of  $\lambda_n$  will transcend  $n^{\text{th}}$ -order causality  $W_n$ , so that the choice of  $\lambda_n$  should be regarded as spontaneous and self-determined at that stage.

In FW so understood, there is libertarian freedom in the sense that the agent's free choice has an inexhaustible causal depth in her cognitive private space. Perhaps, when we scan the inner space of our consciousness, and feel that our choices are free, it is this infinitude that we grasp intuitively, and feel inclined to report as genuine personal freedom.

Identifying human agents with  $F_\infty$  also means that FW is at least a *supertask* [8], a process that involves a sequence of countably infinite number of steps executed in finite time. In Section 5, we will present an argument suggesting that free choice is probably even a *hypertask*, which involves uncountably many steps executed in finite time.

## FREE WILL AS AN INFINITE METATHEORETIC RECURSION

The free agent  $F_\infty$  in our model is an infinite-stage entity straddling the physical  $L_0$  and the “final” or “infinite-th” stage, denoted  $L_\infty$ . For simplicity, we will refer to the agents  $L_0$ -aspect as the “physical aspect” and the  $L_\infty$ -aspect as “transfinite aspect”. In a conflict situation, a response is initiated at the transfinite aspect and transmitted to the physical aspect, where it manifests as the choice  $x_0$ . It stands to reason that the final desination to which information about the sensory input is taken, before the agent’s response is initiated, must also be the transfinite aspect.

### AN INFINITE, STAGED CAUSATION

Extending (7), we can represent the choice  $x_0$  through a sequence of downward causations starting from the “transfinite preference”  $P_\infty$ :

$$x_0 = [[[\dots[[[\lambda_\infty(P_\infty)]\dots]](P_2)](P_1)](P_0) \equiv \lambda_n^\# [P_n, P_{n-1}, \dots, P_\infty](P_{n-1}, \dots, P_0). \quad (8)$$

The interpretation is that  $\lambda_n^\#$ , the will at stage  $L_n$ , is fixed by higher-order preferences, and then selects an outcome  $x_0$  depending on lower-order preferences.

The form of Eq. (8) suggests that the larger the recursion depth  $n$ , the fewer the higher-order preferences  $P_{n+1}, P_{n+2}, \dots$  that could sway  $\lambda_n^\#$  from the motivation encoded by  $P_\infty$ . In Eq. (8), suppose that  $0_{n-1}, \dots, 0_0$  represent the probability functions  $P_{n-1}, \dots, P_0$  that are *unbiased* in the sense of being consistent with  $P_\infty$ .

Replacing the lower-order preferences by their unbiased values in Eq. (8), we now define the  $n^{\text{th}}$ -order *intent*

$$x_n = \lambda_n^\# [P_n, P_{n+1}, \dots, P_\infty](0_{n-1}, \dots, 0_0), \quad (9)$$

meaning that  $x_n$  is the choice  $x_0$  that *would* be made if there are no distortions downwards from level  $L_{n-1}$ . Thus we may call  $x_\infty$  as the “prime intent” or “transfinite intent”, the option that would be selected if the will at infinity,  $\lambda_\infty$ , were to act unthwarted on the physical.

During the act of free choice, the prime intent is replaced stage-wise by lowerorder intents, until the final choice is reached. We may refer to this infinite train

$$\mathbf{X} \equiv x_\infty \rightarrow \dots \rightarrow x_n \rightarrow x_{n-1} \rightarrow \dots \rightarrow x_1 \rightarrow x_0, \quad (10)$$

as the “descent of the will”. This immediately evokes the notion of FW as the effectiveness of communication of  $x_\infty$  from the transfinite aspect to the physical aspect, undistorted by lower-order preferences. The will is stronger, if this channel of communication (“volition channel”) is clearer, uncluttered by lower-order motivations, beliefs and desires inconsistent with their transfinite counterparts. This line of thought forms the basis of the compatibilist FW introduced further in the text.

### RATIONALIST-COMPATIBILIST FREE WILL

The model above can be extended to protect FW from what may be called the ‘rationalist/robot paradox’. By definition, a free, rational agent will, when faced with a choice, select the optimal option. His behavior is completely predictable, assuming that there is a single rational option. For a libertarian, rationality appears to undermine freedom [9]. Now this is not the case, as viewed by a compatibilist. But the rationalist/robot paradox asks how the compatibilist would differentiate a rational agent from an optimal robot programmed to choose rationally.

Following the line of thought indicated previously, we would like to think of the correlation between the prime intent and final choice as a measure of FW, since it expresses how well the agent is able to hold on to her prime intent by overcoming deviating influences. However,

this correlation would stay as merely incidental, unless the physical aspect holds  $x_\infty$  as her purpose or motive. For this, she must have cognizance of  $x_\infty$ . Precisely this defines the role played by the agent’s rational faculty or reasoning in FW. We will refer to an  $F_\infty$  agent equipped with the rational faculty by  $\#F_\infty$ .

The role played by reason is crucial. Without it, the physical aspect has no motive to deviate behavior from that determined by physical causality  $W_0$ . Now there may be random deviations from  $P_0$  (applicable to animal agents), but they would be devoid of any systematic or deliberate attempt to transcend  $W_0$ . By contrast, a human agent, on the recommendation of reason, tries to overcome the imposition of  $W_0$  by trying to deviate  $X_0$  towards  $X_\infty$ . Here the quantity  $X_j$  represents the random variable corresponding to  $x_j$ , i.e., values of variables associated with a probability distribution. Thus the reasoning faculty serves as the basis through which the opportunity provided by causal openness is exploited.

An agent is free to the extent that she is able to enforce her transfinite will on her physical choice. (Complications arising from the corruption of the rational faculty will be ignored here.) This gives us a quantification of Rationalist-compatibilist FW:

$$G = \text{Corr}(X_\infty: X_0), \tag{11}$$

where Corr is any measure normalized so that  $-1 \leq G \leq +1$ . The rational free-willed agent is characterized by  $G = 1$ , while a person completely under the sway of material nature, by  $G = -1$ .

Lacking (substantial) reasoning, an animal may be represented simply by  $F_\infty$ . The animal is free, but not free-willed. We express this insight with the expression:

$$\text{Freedom} + \text{reason} = \text{free will} \tag{12}$$

Quantum matter, or in particular quantum AI, which remains under the scope of physical causality, is a  $F_{0.5}$  agent, while a classical robot is a  $F_0$  agent. It is clear how this Rationalist-compatibilist account protects FW from rationalist/robot paradox: a deterministic robot is a  $F_0$  agent, while a rational free-willed agent is a  $\#F_\infty$  agent with  $G = 1$ .

The brain is arguably a special organ, whose physical structure has somehow been evolved equipping it with a FW oracle, providing a gateway to the transfinite aspect. AI lacks this and the physical laws governing its dynamics are causally closed.

It seems to be an interesting proposition that plants and “lower animals” (like microbes), which lack a central nervous system (CNS), could be considered as intermediary agents between quantum matter and higher animals (like mammals, reptiles and birds, which have a CNS), and thus represented by  $F_K$ , where  $0,5 < K < \infty$ . Some of these ideas are summarized in Table 1.

**Table 1.** Freedom gives spontaneity, reason gives self-determination, and freedom with reason is free will. AI, lacking a FW oracle and being thus just a special configuration of quantum matter, is described by first-order indeterminism. By contrast, the higher animal or human brain, being equipped with a FW oracle, has freedom at all orders. Perhaps the  $F_\infty$  structure, common to humans and animals, is necessary for emotional behavior.

Entity	Agent type	Resource
Human	$\#F_\infty$	Free will (Freedom at all orders, plus reason)
Higher Animals (having a CNS)	$F_\infty$	Freedom at all orders
Lower Animals and plants (Lacking a CNS)	$F_K$ ( $0,5 < K < \infty$ )	Freedom up to a finite order
Quantum AI	$F_{0,5}$	First-order freedom-without-will
Classical AI	$F_0$	Determinism

## CAUSAL VS. LOGICAL DETERMINISM

The resolution of the rationalist/robot paradox shows that the predictability of behavior does not imply that the behavior was causally determined (as in the robot's case). There is logical determinateness about the rational agent's behavior even though he transcends physical causality. We express this idea by:

$$\text{Logical determinism} \Rightarrow \text{Causal determinism.} \quad (13)$$

Now this result would be undermined if all humans were perfect ( $G = 1$ ) then, even if  $P_\infty \neq P_0$ , we would be led to suspect that there is a "law of goodness", characterized by  $P_\infty$ , that controls human behavior. However, some people are imperfect (having  $G < 1$ ), suggesting that human behavior in general transcends causal determinism, and weakening the need to undermine the above conclusion.

## EXPERIMENTAL TEST

It is an interesting and old<sup>1</sup> question: how to experimentally demonstrate the existence of FW? Our model suggests that unfocused or casual acts of choice would be governed by a probability distribution  $P_0$ , while a free-willed action with deliberate intent will in general produce a deviation from  $P_0$  towards  $P_\infty$ .

The observation of discrepancy between the statistics of focused and unfocused choice, could be one way to demonstrate the existence of FW. Designing such an experiment may not be easy, since the very act of focusing may psychologically alter  $P_0$ , so that an observed deviation may either be due to will-induced deviation or due to an alteration of  $P_0$  or due to both.

## WHY WE ARE NATURAL LIBERTARIANS

The model also helps make sense of people's instinctive inclination to Libertarianism. It is a reasonable assumption that as an agent introspectively scans the inner space of her consciousness, depending on how subtle her awareness is, she can at best objectively perceive only so deep as there is freedom.

The unfreedom and higher-order causal influences lying beyond that point become part of her subjective consciousness, and she is unable to consciously experience them, though she may deduce them by observing her conscious choices and preferences.

## FREE CHOICE AND UNCOMPUTABILITY

We now explain a line of thought indicated in Ref. [6], on the correspondence between the causal hierarchy and the hierarchy of Turing uncomputability, in reversed ordering. That is, given a  $\#F_\infty$  agent,  $x_\infty$  is computable, while  $x_m$  is harder to compute than  $x_n$  if  $m < n$ .

The basic idea behind this claim is the following. Suppose one has a computer program (or Turing machine) so powerful that it can compute the free choice of a free-willed agent, using the current most detailed description of her brain state. Now if its prediction were shown to the agent, being free-willed, she may contradict the prediction. The conclusion is that such a powerful computer program does not exist. We now consider a somewhat more detailed argument.

Given a free-willed  $\#F_\infty$  agent  $\mathbf{A}$ , suppose there is a computer  $\mathbf{T}_C$ , programmable in some computer language  $\mathbf{T}_L$  and suitable for the task of computing  $\mathbf{A}$ 's free choice  $x_0$ . Let  $\hat{\mathbf{A}}$  denote the description of  $\mathbf{A}$  as a computer program in  $\mathbf{T}_L$ . We assume that all computer programs that encode in  $\mathbf{T}_L$  the description of free-willed agents are denumerable and that  $\hat{\mathbf{A}}$  is the  $a^{\text{th}}$  program. Similarly, one is assumed to be able to encode situation any conflict situation  $\mathbf{J}$  in the medium of  $\mathbf{T}_L$  by a description  $\hat{\mathbf{J}}$ , and enumerate them alphabetically as some number  $j$ .

If free choice is computable, then the computer  $\mathbf{T}_C$  can, given the enumerations for the  $\mathbf{T}_L$ -description of the agent and the conflict situation, compute the agent's free choice in finite time, or:

$$\tau_C(a; j) = \begin{cases} 0 \Leftrightarrow \mathbf{A}(\mathbf{J}) = 0, \\ 1 \Leftrightarrow \mathbf{A}(\mathbf{J}) = 1. \end{cases} \quad (14)$$

where for simplicity we have assumed the outcome to be two-valued. (There is no loss of generality, since any computable output can be made binary, for example by assigning "0" if the outcome is non-numerical or numerical and less than 0, and "1" otherwise). In words, the computer produces output 1 (resp., 0) if  $\mathbf{A}$  freely chooses 1 (resp., 0) when faced with conflict situation  $\mathbf{J}$ .

For any positive integer  $j$ , using this as a subroutine, we can build another program:

$$\tau_R(j) = \begin{cases} 1 \Leftrightarrow \tau_C(j; j) = 0, \\ 0 \Leftrightarrow \tau_C(j; j) = 1. \end{cases} \quad (15)$$

which is a representation of the above notion of the uncooperative free-willed agent. Thus  $\mathbf{T}_R$  outputs "0" on input  $j$  iff the  $j^{\text{th}}$  free-willed agent outputs "1" on the  $j^{\text{th}}$  conflict situation.

We can now apply the computer to  $\mathbf{T}_R$ , so that from Eqs. (14) and (15):

$$\tau_C(t_R; j) = \begin{cases} 1 \Leftrightarrow \tau_C(j; j) = 0, \\ 0 \Leftrightarrow \tau_C(j; j) = 1. \end{cases} \quad (16)$$

where  $t_R$  is the enumeration of  $\mathbf{T}_R$ . If we set  $j := t_R$ , then we are led to a contradiction as it would entail that  $\mathbf{T}_C(t_R, t_R) = 0$  if and only if  $\mathbf{T}_C(t_R, t_R) = 1$ .

To restore consistency, we infer that  $\mathbf{T}_C$  will never halt on inputs  $(t_R, t_R)$ , which thereby constitutes an undecidable Gödel sentence under the above encoding. We conclude that in general  $x_0$  will be uncomputable for the family of computer programs considered.

One can conceive a higher-level "meta-computer" program  $\mathbf{T}_C[(1)]$  that is able to decide whether  $\mathbf{T}_R(t_R)$  equals 0 or 1, but that is not contained in this family. If the cardinality of such meta-computers is greater than  $\aleph_0$  (countable infinity), then the above diagonal argument based paradox can be averted, because the meta-computers will not be denumerable. This situation is similar to that pertaining to the concept truth, of requiring a metalanguage in order to define truth in the object language [3].

The proof given above for the uncomputability of  $x_0$  is similar to the that of the uncomputability of the halting problem for Turing machines [2]. The concept of a meta-computer alluded to above indicates that the proof of uncomputability "relativizes", meaning that one can construct harder problems, by allowing a computer program to call as subroutine an "oracle" that solves the above free choice problem in bounded time. One can then construct a Gödel sentence for this oracle-enhanced program, which then yields a problem with its hardness shifted one level higher than the free choice problem above. Problems on the same level of uncomputability, i.e., uncomputable problems which are Turing-equivalent, form a Turing degree. The process can be repeated to construct higher Turing degrees, i.e., the next higher levels of more uncomputable problem [10]. It is known that there are  $2^{\aleph_0}$  (uncountably many) Turing degrees.

We suggest that the the causal hierarchy corresponds to Turing degrees, but in inverse ordering, whereby the prime intent, which arises beyond an infinite number of causal stages, is itself Turing-computable. But the descent of the will would correspond to transition to higher levels of Turing uncomputability, making the free choice of agents of sufficiently low FW highly uncomputable. The act of FW in general must be a monstrous hypertask, since the lower causal stages correspond to ever higher orders of uncomputability.

Why should the causal and computational hierarchy correspond with each other? Here we will appeal to a teleological argument: that if the consequences of the causal hierarchy were computable, then there would have been no need for the “brute force” computation provided by the physical manifestation of the universe and human agents!

Considerable research has been devoted in computability theory and mathematical logic to the study of the complicated structure of Turing degrees. Perhaps all of that may have a bearing on the cognitive structure of free-willed agents.

## **NEUROLOGICAL BASIS FOR FW**

The presence of the FW oracle in the human brain marks the basic difference between a human agent and a robotic simulation. The question of how the FW oracle is embedded in the brain and called forth, is briefly considered here.

A  $\#F_\infty$  agent is an infinite entity, whereas the physical brain of a human being is finite in terms of its information storage and computation capacity. Therefore, if humans are  $\#F_\infty$  agents, sufficiently high levels  $L_j$  cannot have a physical representation, i.e., a neural correlate.

We propose the following physical realization of the model. At the instance a human becomes aware of a choice, she is instinctively driven to enact her “nature”, encoded by  $P_0$ . At the selection window, her “reason”, which carries a representation of  $P_\infty$ , advises her to deviate  $X_0$  towards  $X_\infty$ . This creates a potential tension, which may result in a fleeting quantum superposition, and may correlate with the agent subjectively experiencing an internal conflict.

Since  $P_0$  is determined by  $W_0$ , its neural correlate is expected to be well defined, and associated with the motor cortex. Since  $P_\infty$  is largely computable, its neural correlate is also expected to be well defined, and associated with the reasoning circuits in the pre-frontal cortex. In Figure 1, these two correlates are represented as the slow and fast neural pathways, at whose confluence the FW oracle lies.

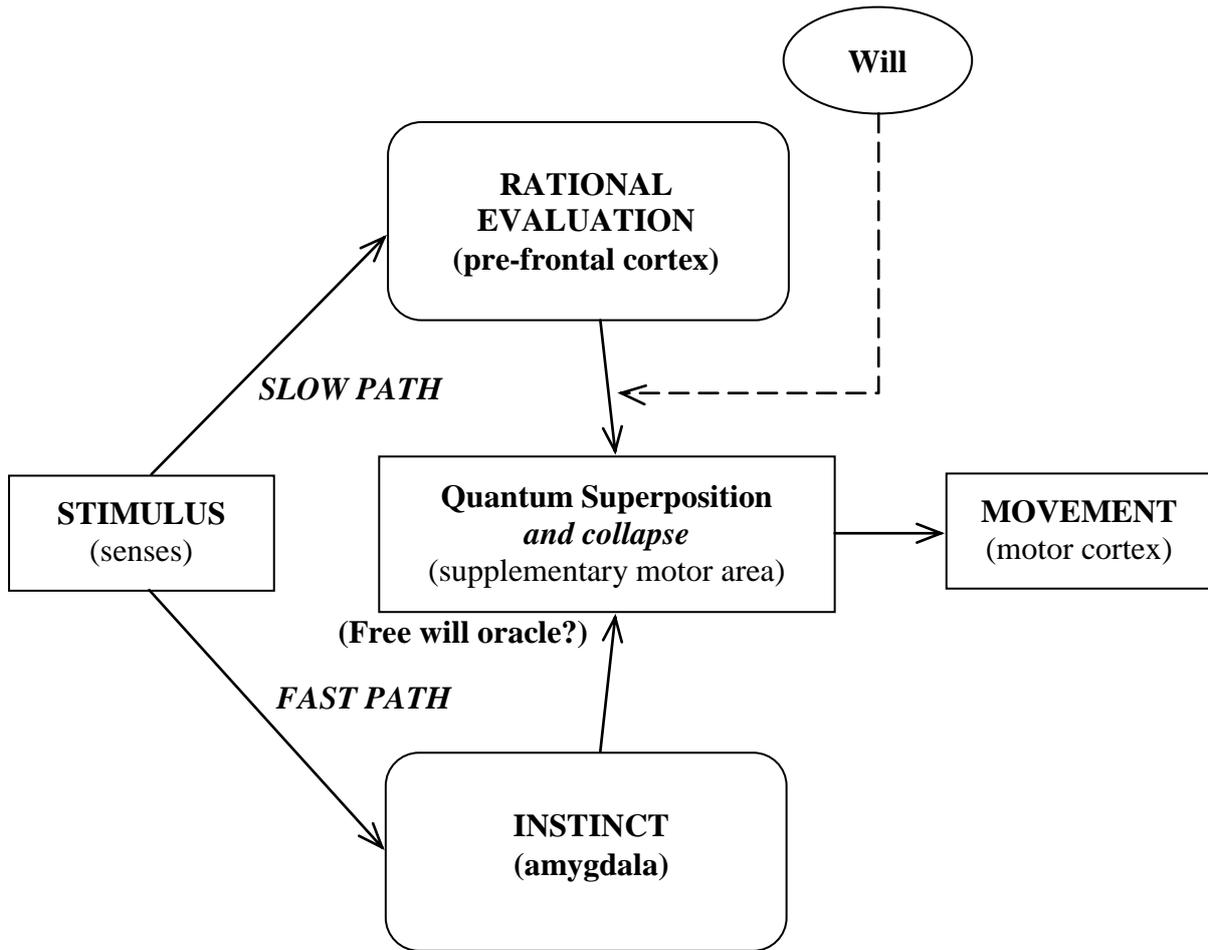
From physically observable data, one may be able to predict a pattern of behavior. However, since the higher levels in the causal hierarchy are not observable, therefore in any given instance of an agent’s free choice, even the most detailed neural imaging (say via fMRI) will be unable in principle to predict with full certainty what the agent will select.

The FW mechanism in the brain of a  $\#F_\infty$  agent can never be modelled as a finite input-output device. An important physical consequence is that, the mental causation that produces a deviation from  $W_0$  may correspond to deviations from physical laws associated with  $W_0$ , like energy conservation or the Second Law of thermodynamics. For example, the initial voltage fluctuation in a motor neuron that initiates the spontaneous movement of a mouse’s whisker, may be energy-wise unaccounted physically, even though the subsequent nonlinear amplification of that fluctuation to a physical action will certainly be governed by (classical) physics.

A possible experimental test of the model could aim to distinguish willful from casual choice. In the latter case, the probability of choice will reflect  $P_0$ , which can be estimated from the relative strengths of signals, as picked up by fMRI scans. Under willful choice, there will be a deviation in the probability of choice away from  $P_0$  towards  $P_\infty$ .

## **DISCUSSIONS AND CONCLUSIONS**

We showed that, although metaphysical libertarianism is provisionally protected against the dilemma of determinism by the two-stage model, it is vulnerable to the recursive version of the dilemma. As a defense for libertarian freedom, we proposed the infinitely recursive two-stage scheme. A free agent is described here by  $F_\infty$ , and free-willed agents by  $\#F_\infty$ . The concept of FW here is of the Rationalist-compatibilist kind.



**Figure 1.** Neurobiological flowchart for free choice by an agent. The conflict situation triggers neural signals along two paths: the fast “nature” path and the slow “reason” path-carrying possibly opposing recommendations for action. If the neural signal in the slow path is weak, the agent executes the instinctive action induced by the fast path. But if the signal in the slow path is sufficiently strong, and there is a conflict between  $P_0$  and  $P_\infty$ , then a quantum superposition is set up at the FW oracle, as the intermediate step for deviating  $X_0$  towards  $X_\infty$ .

Some other issues, with ramifications for quantum physics, neuroscience, mathematics, philosophy, computation theory, are briefly mentioned below.

There appears to be a parallelism between the will in  $F_1$  and hidden variables in ontological models of quantum mechanics [6]. But there are two basic differences: these ontological models attempt *explain* probabilistic physical laws  $W_0$  (and thus correspond to  $F_{0.5}$ ), whereas the will in  $F_1$  may produce *deviations* from  $W_0$ . Second, the will in  $F_1$  is the first rung in an infinite hierarchy of higher-order willings, whereas hidden variable models of quantum mechanics stop at unit depth.

Neurobiology is the area most affected by our model, and also its possible clearest testing ground. Experimentally locating the seat of the FW oracle in the brain, and working out how mental causation initiates free-willed action in motor neurons will be vital. Experiments that distinguish willful and casual choice offer another window of study the neurobiological circuitry for FW.

This understanding can be medically useful. By potentially clarifying the roles neurotransmitters or receptors play in the process of free choice, it may be able to suggest

medical solutions that help encourage self-controlled behavior, by enhancing the “reason” pathway, rather than momentarily suppressing the “nature” pathway (Figure 1). Such treatments may be useful for patients suffering from neuropathological ailments like obsessive-compulsive disorder (OCD).

In mathematics and computation theory, the relationship between the causal hierarchy and Turing degrees would merit further study. This will help to elucidate the scope of AI. The formalization of FW as presented here, along the lines of formalization of the concept of truth [3] would be the first step here.

Our model implies that high FW correlates with improved predictability, i.e., reduced entropy in  $X_0$ . This reduction comes not by a compensatory increase in entropy elsewhere in the universe as required by the Second Law of thermodynamics, but by means of the deviation from  $W_0$  produced by mental causation. Departure of  $P_j$  from  $0_j$  is expressed as a conflict, and thus entropy, on  $L_0$ . A purely physical means to reduce entropy would be subject to the Second Law, with no implications for the agent’s cognitive freedom. But by freely reducing this entropy, the agent is aligning her  $P_j$ ’s with  $0_j$ ’s, and enhancing her freedom. Thus concepts like moral responsibility and justice are helpful as props that encourage free-willed behavior, and thereby help reduce disorder, though not necessarily on the physical level.

This in turn has implications for evolutionary biology. It suggests that the underlying force driving evolution was perhaps not Nature’s quest for propagating the species most successful at survival, but instead Nature’s quest for greater freedom. Darwin-like incremental evolution evolves quantum matter ( $F_{0,5}$  agent) through lower animals lacking a CNS ( $F_J$  with finite  $J$ ), and then through higher animals equipped with a CNS but no cerebral cortex (free agents,  $F_\infty$ ), and from there, finally to free-willed agents,  $\#F_\infty$ . Thus *Homo sapiens sapiens* perhaps already represents the limits of Darwinian biological evolution. The remaining evolutionary journey, towards greater freedom of the will, is now “up to us”. It should be accomplished through self-determination.

## REMARK

<sup>1</sup>The Royal Norwegian Academy of Sciences (KNVA), the predecessor of DNVA, posed in 1839 the academic question “Is it possible to demonstrate human free will and self consciousness?”

## REFERENCES

- [1] Pink, T.: *Free Will: A Very Short Introduction*. Oxford, 2004, <http://dx.doi.org/10.1093/actrade/9780192853585.001.0001>,
- [2] Davis, M.: *Computability and Unsolvability*. Dover, 1982,
- [3] Tarski, A.: *The Semantic Conception of Truth and the Foundation of Semantics*. *Philosophy and Phenomenological Research* **4**(3), 341-376, 1944, <http://dx.doi.org/10.2307/2102968>,
- [4] Heisenberg, M.: *Is free will an illusion?* *Nature* **459**, 164-165, 2009, <http://dx.doi.org/10.1038/459164a>,
- [5] James, W.: *The Dilemma of Determinism: The Will to Believe*. Dover, New York, 1956,
- [6] Nayakar, C.M. and Radhakrishna, S.: *Gödel, tarski and the conundrum of free will*. In: Balakrishnan, J. and Sreekantan, B.V., eds.: *Nature’s Longest Threads*. New Frontiers in the Mathematics and Physics of Information in Biology. World Scientific Publishing Co., 2014,

- [7] Schopenhauer, A.: *On the freedom of Human Will*. 1939, translated from German to English in Zucker, M.: *The Philosophy of American History: The Historical Field Theory*. 1945,
- [8] Laroudugoitia, J.P.: *Supertask*. In: Zalta, E.N., ed.: *Stanford Encyclopedia of Philosophy*, 2009,
- [9] Pećnjak. D.: *Complex Freedom*. *Interdisciplinary Description of Complex Systems* 7(1), 14-21, 2009,
- [10] Lerman, M.: *Degrees of unsolvability*. *Perspectives in Mathematical Logic*. Springer-Verlag, Berlin, 1983.

# THE UNIVERSE IS LIKE A HOLLOWED SPHERE. THE WAVE CONCEPT OF TIME

Andrzej W. Głuszak\* and Franciszek Kluza

University of Life Sciences  
Lublin, Poland

DOI: 10.7906/indecs.13.3.3  
Regular article

*Received:* 9 April 2015.  
*Accepted:* 23 June 2015.

## ABSTRACT

There is space for new ideas of the essence and the entity of time. The article refers to our time concept as a special wave type and presents results of our investigations on this subject. Thus, time defined as waves and an energy carrier could give explanation to multiple unclear phenomena. It could explicate gravity, organization in the planetary systems and light speed limit. A hypothesis that matter exists due to time wave motion would emerge from the elementary particle mass generation by the waves. Time becomes the main driving force in the Universe. The discussed thoughts need further analyses and verification but their confirmation may mean civilization changes.

## KEY WORDS

time, wave, matter, light speed

## CLASSIFICATION

JEL: Z10

PACS: 01.55.+b, 04.20.Cv, 04.20. Gz, 04.30. Db

\*Corresponding author, *η*: gluszak.andrzej@wp.pl; +48 69 378 9027;  
ul. Dziewanny 12/12, 20-539 Lublin, Poland

## INTRODUCTION

The issue of time has been a major subject of scientific exploration from the ancient times through philosophers' cogitation until the present day. Einstein with his theories of relativity made substantive contributions to human understanding of the nature of time. This article is based to a great extent on the transformed mass-energy equivalence law.

There exist multiple concepts of time but room for novel ideas is still opened. Encyclopaedia Britannica defines time as "a measured or measurable period, a continuum that lacks spatial dimensions. Time is of philosophical interest and is also the subject of mathematical and scientific investigation" [1].

In antiquity, time was presented in many ways. According to Heraclitus "everything flows and nothing abides" as he considered time as a cyclical continuum [2]. Zeno of Elea developed the idea that time and space are continuous and infinite and thus Achilles can never overtake the tortoise as the philosopher stated in one of his numerous paradoxes [3]. Aristotle, in turn, defined time as "the number of movement in respect of *before* and *after*" [4].

Saint Augustine believed that time and motion are associated with material time and its passage. Undoubtedly, for him God was the creator of both [5].

Newton's view of time included the idea of only one universal time as an ever uniform stream that was constant and immutable. He stated that time and space have an objective and absolute nature [6].

Bergson distinguished between two forms of time, pure time which is real duration and continuity, from mathematical time which is divisible into units [7].

Kant claimed that time and space are a priori forms of inner senses existing in the human mind which is the subjective source of them [8].

In Husserl's view, time and space were the forms which can be understood using the method of phenomenology. He perceived knowledge as the foundation for synthesis of objects in time [9].

For Leibniz, time and space were only illusions [10]. Yang proposed a quantum theory in which time is a quantum variable with a discrete spectrum, and which is nevertheless consistent with special relativity [11].

According to Tejman, energy (energetic matter), space and time are one. Wave theory introduces energetic matter – a single inflationary force – as the main creation of nature. Energetic matter creates wave formations and they, in turn, create everything [12].

In classical physics, time is an autonomous value independent of others that flows at the same rate in the entire Universe.

In Einstein's general theory of relativity, time stops being independent but is the fourth dimension of spacetime curved by gravity. Consequently, clocks at higher potentials in the gravitational field run at a different rate, time flows slower. Here, gravity is explained as a result of a "curvature" of spacetime caused by structures, accumulated matter. In the special theory of relativity, Einstein determined that the rate at which time passes depends on the speed of the reference frame where time is measured. In the case of the frames moving at speeds approaching the speed of light, significant differences in time measurements occur [13-15].

## DISCOURSE

What is time in its very nature? Are we observing the consequences of time or the actions taking place in the present? How did it come into being and by what means is it spreading? In what way does energy transmit through the empty space?

We can try to find the answers in the formula  $E = mc^2$ , where:  $E$  – energy,  $m$  – mass,  $c$  – defined as speed of light in vacuum. This fundamental message contains a lot of information. Without motion, there is no energy, because if  $c = 0$ , then  $E = 0$ .  $c$  must be greater than 0, otherwise it would not be motion. The transformation of  $m = E/c^2$  leads to the conclusion that there is no mass without energy. Motion  $\rightarrow$  Energy  $\rightarrow$  Mass. As early as in ancient Greece, Anaximander observed that motion is inherent in matter [16].

After the appropriate transformation of the formula  $E = mc^2$  we obtain:  $t = d\sqrt{m/E}$ , where  $t$  – time,  $d$  – distance (length). In sequence: it is known that for uniform motion we use the formula for speed  $v = d/t$ . Since light travels in uniform motion, the considered here  $v = c = d_c/t_c$ , where  $d_c$  and  $t_c$  are distance and time derived from the speed of light. We can write:  $c = d_c/t_c$ . Let us transform the formula  $c^2 = E/m$  into  $(d_c/t_c)^2 = E/m$ .

Quantity  $c^2$  is always greater than 0 and mass  $m$  is positive in physical terms, which means that  $E > 0$ . We apply root extraction on both sides of the equation, and since  $E$  and  $m$  are greater than 0, we can omit the brackets for absolute values obtaining:  $d_c/t_c = \sqrt{E/m}$ . Finally,  $t_c = d_c\sqrt{m/E}$ . Time  $t$  measured by clocks is modified by gravity [14]. We stated, it is not identical to “pure” time, cosmic time  $t_c$ .

Quantity  $c^2$  in the formula raises doubts and provokes some questions. Whereas the kilometre is a simple continuation of its standard, e.g. the metre, the square kilometre is a surface with an infinite number of shapes. How does the second squared change? What dimensions does  $(\text{km/s})^2$  transform into? This problem can be solved by applying the same units for  $t_c$ ,  $E$ ,  $d_c$  and  $m$ .

Here, Anaximenes with his view on the unity of nature: “the underlying substance was one and infinite” [17] and Anaxagoras, with his “In everything there is a share of everything” [18] should be referred to.

If  $t$  is greater than 0 at present, it means *panta rhei*.  $d > 0$  follows the idea of Parmenides that there is existence but no-existence is not [19]. If  $m > 0$ , this indicates that objects existing in time always have mass, they are not massless. If  $E > 0$ , this implies that currently energy exists everywhere.

In the cosmogonic aspect  $t_c > 0$ ,  $d_c > 0$ , which means that neither time nor space existed before the Beginning. They manifest themselves with energy and elementary particles after the Big Bang explosion as before it the space-time was squeezed to a very small volume [20]. “All things were together” as Anaxagoras put it [18].

Critically analysing current knowledge, guided by Ancient Greek thoughts let’s assume the wave concept of time.

One of the obvious properties of time is its ability to penetrate. It runs through and crosses every boundaries.

Time considered as waves, with the appropriate conformation, could satisfy the conditions for boundless penetration of the Universe. The length of such a wave should be beyond  $c$  values that is over 300 000 kilometres and to square 90 billion km. Whereas frequency may be sought from Planck’s time to infinity.

Time defined as the waves, as the carrier of elementary energy could cause certain implications and explain many unclear phenomena.

Relations between time and energy are found at the Heisenberg’s uncertainty principle. Einstein argued that “Heisenberg’s uncertainty equation implied that the uncertainty in time was related to uncertainty in energy, the product of the two being related to Planck’s constant” [21, 22]. As early as in 1924, Louis de Broglie claimed that every particle of

velocity different from zero may be associated with a wave of defined frequency and length. Besides, he also proposed that all the material objects have a wave nature [23].

The formula  $t = d\sqrt{m/E}$  (for  $d_c$  and  $t_c$ ) implies that if  $E = m$ , then  $t = d$ , that is, time is equal to distance. The junction unifying time and dimensions uncovers here. At this “point-moment”, when  $c = 1$ , a time wave (an active factor) carrying energy generates mass to an elementary particle (a passive element). It means matter exists within and due to time motion.

Time wave  $\rightarrow$  elementary particle clusters  $\rightarrow$  matter center. As Heraclitus wrote, “Everything flows, nothing stands still. The one constant in the whole process is the law of change by which there is an order and sequence to the changes” [2]. Is this law the Formula?

Arguably, the mass-energy equivalence transformation involves chip – a Higgs boson.

It may be assumed that the time shock waves initiated after the Big Bang are composed of n-multi waves of different wave lengths and frequency that are interrelated, overlapping and looping, making strings, beams and streams. And the vacuum appears in the centre of the Space. The Universe is like hollowed sphere.

We put forward that some of the waves penetrate accumulated matter and flow forward, while others separate bounce off and create centripetally rotating loops responsible for the phenomenon of gravity. The strength of this and energetic reactions comes from continuous movement of the oncoming time waves. Time waves  $\rightarrow$  accumulated matter  $\rightarrow$  spacetime curvature  $\rightarrow$  gravity.

If not the Aether than possible it is time stream approaching from the abyss of the Universe and falling under a certain angle, that form elliptical orbits of planets. Similarly to mechanical and electromagnetic impact, the time wave beam running between Earth and Venus could cause the retro-grade rotational motion of the planets [24, 25]. Moon, which is a fragment detached from Earth, exhibits tidal locking [26, 27]. It has been an immanent part of our planet and kept tied by a specific unique gravitational force. We propose this natural satellite is held by the globe as if on ropes. The lunar mascons might make a kind of hook for terrestrial gravity [28].

What be the consequence of time and its energy reversing? We presume if time gathering elementary particles flowed in the opposite direction, it could form Black Holes. It might curl up together with space and matter till it has attained the critical point. Then the Big Bang would occur and the process would start all over again. Let us think about the possibility that Black Holes are a source of new worlds. Answering Stephen Hawking’s question – it is time that breathes fire into equations and makes the Universe for them to describe [29]. We suppose the number of dimensions might depend on amount of directions of time streams.

Why  $c$ , denoting the speed of light in a vacuum, cannot exceed 299 792 458 m/s? [30]. What are the limits? Perhaps time is the barrier, more specifically the oscillations of its waves, and the answer could be the shortest time wave-length ca. 300 000 km, with one oscillation each second. Light beam would stop for an immeasurably short period, perhaps Planck time, i.e. ca.  $5,4 \cdot 10^{-44}$  s, at the oscillation “point-gate” of the time wave [30, 31]. At this point, the theory of time discontinuity is worth mentioning along with a chronon unit [32, 33]. To conclude, light travels jumping at constant velocity.

We suggest very high density of time texture, where discontinuities of the spacetime are extremely small, maybe close to the Planck length, that is the order of  $1,6 \cdot 10^{-35}$  m. However, taking into account the enormous length of waves and possibilities of the current measuring instruments, the time waves remain immeasurable. Planck’s equation for the elementary particle, photon, can be written as  $E = hc/\lambda$  [34]. This means that the longer wave, the lower energy. With a wavelength ( $\lambda$ ) of 299 792 458 m, energy equals to Planck’s constant

( $6,626 \cdot 10^{-34}$  J·s, we call it primal energy). We introduce a concept that the interrelation between light and time is very special. The limitation within the time loop is responsible for the masslessness of the photon.

The Universe expands and time waves could also alter their properties, so together with this metamorphosis the existence will change.

## CONCLUSION

The Universe is like a hollowed sphere. Time is the main driving force in the Cosmos. The whole energy of space comes from the Big Bang and is spread and carried by time waves.

Time treated as the waves could explicate such phenomena as matter existence and gravity. Oscillations of time waves limit light speed and are responsible for masslessness of the photon. The nature of time waves makes their detection unable using the present research methods.

The results of our basic research presented in the article are to signalize the need for putting forward questions concerning the essence of time. The responses may open up a new chapter in science and lead to civilization changes. Currently, we do not influence time.

Will it ever change? Only time will tell.

## REFERENCES

- [1] –: *Encyclopaedia Britannica*.  
<http://www.britannica.com>, 2012,
- [2] Kahn, C.: *The Art and Thought of Heraclitus: Fragments with Translation and Commentary*. Cambridge University Press, London, pp.1-23, 1979,
- [3] Huggett, N.: *Zeno's Paradoxes*.  
Stanford Encyclopedia of Philosophy, 2010,
- [4] Irwin, T. and Fine, G.: *Aristotle; Selections*.  
Translated with Introduction, Notes, and Glossary. Indianapolis, Hackett, 1995,
- [5] O'Daly, G.: *Augustine's Philosophy of the Mind*.  
University of California Press, Berkeley, 1987,
- [6] Westfall, R.S.: *Isaac Newton*.  
Cambridge University Press, Cambridge, 2007,
- [7] Guerlac, S.: *Thinking in Time: An Introduction to Henri Bergson*.  
Cornell University Press, Ithaca, 2006,
- [8] Rohlf, M.: *Immanuel Kant*.  
In: Zalta, E.N., ed.: *The Stanford Encyclopedia of Philosophy*. 2010,  
<http://plato.stanford.edu/archives/fall2010/entries/kant>,
- [9] Derrida, J.: *The Problem of Genesis in Husserl's Philosophy*.  
University of Chicago Press, Chicago and London, 2003,
- [10] Leibniz, G.W.: *Philosophical Texts*.  
Translated and edited by Woolhouse, R.S. and Francks, R. Oxford University Press, Oxford, 1998,
- [11] Yang, C.N.: *On quantized space-time*.  
*Physical Review* **72**, 874, 1947,
- [12] Tejman, C.: *United Nature and Theory*. Book 1: *Grand Unified Theory: Wave Theory*.  
Chapter 8., 2001,  
<http://www.grandunifiedtheory.org.il>,
- [13] Hartle, J.B.: *Gravity: an Introduction to Einstein's General Relativity*.  
Addison-Wesley, San Francisco, 2003,
- [14] Mermin, N.D.: *It's About Time. Understanding Einstein's Relativity*.  
Princeton University Press, Princeton, 2005,

- [15] Rindler, W.: *Relativity. Special, General and Cosmological*.  
Oxford University Press, Oxford, 2001,
- [16] Couprie, D.L.; Hahn, R. and Naddaf, G.: *Anaximander in Context: New Studies in the Origin of Greek Philosophy*.  
York Press & State University of New York, Albany, 2003,
- [17] Kirk, G.S.; Raven, J.E. and Schonfield, M.: *The Presocratic Philosophers*. 2<sup>nd</sup> edition.  
Cambridge University Press, Cambridge, 1983,
- [18] Curd, P.: *Anaxagoras*.  
In: Zalta, E.N., ed.: *The Stanford Encyclopedia of Philosophy*, 2011,  
<http://plato.stanford.edu/archives/fall2011/entries/anaxagoras>,
- [19] Melchert, N.: *The Great Conversation: A Historical Introduction to Philosophy*.  
McGraw Hill, Boston, 2002,
- [20] Whitney, C.: *NASA Technology Views Birth of the Universe*.  
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, 2014,
- [21] Heisenberg, W. *Physics and Philosophy: The Revolution in Modern Science*.  
Harper & Row, New York, 1962,
- [22] Gamow, G.: *The great physicist from Galileo to Einstein*.  
Harper Brothers, 1961 & Dover, 1988,
- [23] de Broglie, L.: *The Reinterpretation of Wave Mechanics*.  
Foundations of Physics **1**(1), 5-15, 1970,
- [24] Bakich, M.E.: *The Cambridge planetary handbook*.  
Cambridge University Press, Cambridge, 2000,
- [25] Correia, A.C.M.; Laskar, J. and Néron de Surgy, O.: *Long-term evolution of the spin of Venus. I. theory*.  
Icarus **163**(1), 1-23, 2003,  
[http://dx.doi.org/10.1016/S0019-1035\(03\)00042-3](http://dx.doi.org/10.1016/S0019-1035(03)00042-3),
- [26] Alexander, M.E.: *The Weak Friction Approximation and Tidal Evolution in Close Binary Systems*.  
Astrophysics and Space Science **23**(2), 459-508, 1973,  
<http://dx.doi.org/10.1007/BF00645172>,
- [27] Zhang, J. et al.: *The proto-Earth as a significant source of lunar material*.  
Nature Geoscience, Advance Online Publication, March 28, 2012,
- [28] Jolliff, B.L.; Wieczorek, M.A.; Shearer, C.K. and Neal, C.R.: *New Views of the Moon*.  
Reviews in Mineralogy & Geochemistry **60**, Plate 3.1, 2006,
- [29] Hawking, S.: *A Brief History of Time*.  
Bantam Books – Bantam Dell Publishing Group, 1988,
- [30] Taylor, B.N. Mohr, P.J., eds.: *The NIST Reference on Constants, Units, and Uncertainty*.  
Fundamental Constants Data Center, NIST Physical Measurement Laboratory, 2010,  
<http://physics.nist.gov/cuu>,
- [31] Lesch, H. and Zaun, H.: *The Shortest Story of All Life*. In German.  
Piper-Verlag, Munich, 2008,
- [32] Albanese, C. and Lawi, S.: *Time Quantization and q-deformations*.  
Journal of Physics A: Mathematics and General **37**(8), 2983-2987, 2004,  
<http://dx.doi.org/10.1088/0305-4470/37/8/009>,
- [33] Caldirola, P.: *The Introduction of the Chronon in the Electron Theory and a Charged Lepton Mass Formula*.  
Lettere al Nuovo Cimento **27**(8), 225-228, 1980,  
<http://dx.doi.org/10.1007/BF02750348>,
- [34] Alonso, M. and Finn, E.J.: *Fundamental University Physics*. Volume III: *Quantum and Statistical Physics*.  
Addison-Wesley, 1968.

# THE ART OF MEMORY AND THE GROWTH OF THE SCIENTIFIC METHOD\*

Gopal P. Sarma\*\*

Emory University  
Atlanta, U.S.A.

DOI: 10.7906/indecs.13.3.4  
Regular article

Received: 28 March 2015.  
Accepted: 23 June 2015.

## ABSTRACT

I argue that European schools of thought on memory and memorization were critical in enabling growth of the scientific method. After giving a historical overview of the development of the memory arts from ancient Greece through 17<sup>th</sup> century Europe, I describe how the Baconian viewpoint on the scientific method was fundamentally part of a culture and a broader dialogue that conceived of memorization as a foundational methodology for structuring knowledge and for developing symbolic means for representing scientific concepts. The principal figures of this intense and rapidly evolving intellectual milieu included some of the leading thinkers traditionally associated with the scientific revolution; among others, Francis Bacon, Renes Descartes, and Gottfried Leibniz. I close by examining the acceleration of mathematical thought in light of the art of memory and its role in 17<sup>th</sup> century philosophy, and in particular, Leibniz's project to develop a universal calculus.

## KEY WORDS

scientific method, scientific revolution, the Enlightenment, methodological thinking, universal calculus

## CLASSIFICATION

JEL: B19, O31

PACS: 01.

\*The title is adapted from Chapter 10 of Frances Yates' book *The Art of Memory*.

\*\*Corresponding author, *✉*: gopalsarma@outlook.com; +1 (404) 727-6123;  
201 Dowman Drive, Atlanta, Georgia 30322 USA

## INTRODUCTION

“What is the scientific method?” It is a question that rarely makes an appearance in the scientific world. And perhaps for good reason. Most scientific disciplines are sufficiently advanced that to the extent that there is any “method” for students to learn, it is largely implicit knowledge that is absorbed by actively problem solving and participating in the frontiers of research.

According to the Oxford-English Dictionary, the scientific method is “a method of procedure that has characterized natural science since the 17<sup>th</sup> century, consisting in systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses” [1]. There is no doubt that good science is characterized by the qualities encapsulated in the above definition, and as I stated above, that these qualities are learnt implicitly by participating in the scientific process under the guidance of more experienced researchers. And yet, for an idea that is so simply stated, there seems to be a profound sense of mystery surrounding the historical events that allowed for the scientific method to take root at an institutional level. What was it that took place in Europe during the 16<sup>th</sup> and 17<sup>th</sup> centuries that gave rise to a cultural transformation so significant that it changed the face of not only one area of systematic inquiry, but the entirety of the human pursuit of knowledge? And how did a philosophical transformation of massive proportions happen during this period and why did it not happen before? The world had certainly seen systematic scientific reasoning and hypothesis-driven investigation prior to the 16<sup>th</sup> and 17<sup>th</sup> centuries. Empires had existed for thousands of years, and the very operational foundation of these multinational entities would have required advanced knowledge to maintain such vast infrastructure. We can point to numerous examples of sophisticated understanding in cosmology, medicine, or mathematics that are clear evidence of some form of principled reasoning that existed prior to what we now call the scientific revolution.

And yet, it does seem that something profound happened during this time period that deserves explanation. As modern intellectuals, we are surrounded by numerous forms of print and electronic media that allow us to participate in a philosophically reflectively culture that might shape our world views gradually. We can read and re-read articles in newspapers, scientific journals, books, and magazines, we can watch and re-watch lectures on the Internet, and follow up these periods of intense investigation with rapid electronic discussion with colleagues or anonymous commentators on blogs and question answering sites. And to nurture “pre-scientific” or “meta-scientific” knowledge, elite academic institutions have created exclusive departments for philosophy and history of science.

Almost none of these outlets and intellectual infrastructure existed in the 16<sup>th</sup> and 17<sup>th</sup> centuries, and certainly not in the period leading up to it. What then would have enabled a highly philosophical stance toward the pursuit of knowledge, a “meta-idea,” to take root? Even in the modern world, philosophically minded young people are chastened by their more worldly and grounded peers to focus on concrete problems. Theoretical physicists lament the towers of abstraction built by mathematicians and engineers often hold a similarly derisive attitude towards their colleagues in physics. In a world without all of these distinctions, what would have been the motivating forces for adopting a philosophical and abstract perspective with regards to the rational pursuit of knowledge, and not simply in the hands of a few brilliant visionaries but rather at an institutional and cultural level? Why did these perspectives gain momentum and why had they not before? The profoundly significant differences between our own world – even the recent world prior to the Internet – and European intellectual climate several centuries ago suggests that perhaps there were important factors that have been forgotten.

Why should we care to answer these questions? Other than the intrinsic historical value, there are perhaps more pressing contemporary concerns. We live in an age in which the scope of science is larger than it has ever been, in which more scientists are being trained than ever before, and in which the tectonic plates of the scientific establishment are shifting, with such headlining trends such as the massive, multi-institution, multi-national “big science” projects, or the petabytes upon petabytes of data being produced which demand a global computational infrastructure unprecedented in scope. In response to these trends, many commentators have argued that we are entering a new phase of science entirely [2] and some have even questioned the role hypothesis driven investigation will continue to play in the era of “big data” [3]. There may or may not be lessons for us, but in such an era of turbulence and upheaval, it seems worth re-examining the origins of a previous scientific revolution.

My perspective is that of a practicing scientist who has always been mystified by the circumstances surrounding the scientific revolution. Contemporary scientists are constantly challenged by the problem of how to convince others to pay attention to their work. In some areas, such as computer science, practical industrial applications often drive interest and funding, while in others, such as in theoretical physics, elegance, simplicity, and vision can often be the key forces. New ideas are constantly filtered through the attentional lenses and biases of a very human institution. How then did the basic, philosophical notions of the scientific revolution take root and why had they not done so earlier? Were practical concerns key factors in influencing adoption of a scientific outlook? Or were inspiration and vision? Or some combination of both?

While I hope that the narrative I have stitched together here is a compelling and coherent story, it is not intended as a complete one. I have made the decision to tell the story of the scientific method from the perspective of the art of memory, and leave out the many other important factors which are more widely known. It goes without saying that a proper treatment would attempt to integrate as many different social, intellectual, and psychological factors as possible. Having stated these caveats, I believe that this article may serve as a useful starting point for those in a position to investigate these ideas more fully.

The position that I advocate in this article is that the scientific revolution was indeed a period of dramatic intellectual change and that in a sense, what took place was a philosophical transformation. But it is for reasons that I suspect most people will be wholly unfamiliar with. Drawing primarily from the work of Frances Yates and Paolo Rossi [4, 5], I will argue the following:

- During the 16<sup>th</sup> and 17<sup>th</sup> centuries, Europe was a cauldron of ideas related to the art of memory,
- These ideas have their origin in ancient Greece and Rome, where the practice of memorization, using a visualization technique known as the “method of places and images,” was a foundational methodology for rhetoric.
- The classical art of memory was preserved through the Middle Ages largely in the monastic context, with St. Thomas Aquinas being one of the principal champions of the method of places and images. Aquinas also introduced an important innovation by suggesting that these techniques could be used not only for the purpose of memory, but also as a concentration device for developing virtue and ethical behavior.
- During the Renaissance, a critical transition took place wherein the mnemonic images of the classical art of memory were thought not only to serve the purpose of remembering, but also to represent the logical structure of nature itself.
- In the 16th century, Giulio Camillo, a major intellectual figure in the Renaissance transformation of medieval memory, laid out an ambitious program to systematize all knowledge with a theatre sized repository of standardized mnemonics.

- Somewhat independent of the classical art of memory, the 13th century Franciscan friar Ramon Lull proposed an alternate conceptualization of memory techniques in the context of combinatory wheels whose purpose was to eliminate logical contradictions from Christian theology.
- At the arrival of the 17<sup>th</sup> century, there were several major schools of thought on memory and method in place. The classical art of memory involving the method of places and images, the method of Ramon Lull and his combinatory wheels, and finally, the dialectic method of Petrus Ramus, with its emphasis on a logical structuring of knowledge and memorization by ordinary repetition.
- In the late 16<sup>th</sup> century and 17<sup>th</sup> centuries, Giordano Bruno and Gottfried Leibniz emerged as significant unifiers of these disparate and often conflicting intellectual traditions. Bruno developed a hybrid mnemonic-Lullian method in which the method of places and images was used in conjunction with Lullian combinatory wheels, and Leibniz developed a hybrid mnemonic-dialectic method in which an encyclopaedia was first developed to which each concept would be assigned a mnemonic image.
- The fundamental concept in all of these developments is the general notion of “method,” as first proposed by Petrus Ramus. The connotation of “method” is likely analogous to our modern word algorithm, and in the same way that the true meaning of algorithm only gains in substance when used in the context of computing technology, the word method gained its novel connotation in the context of the method of places and images, Lullian combinatory wheels, and the subsequent advances made by Bruno and Leibniz.
- It was amidst these developments in methodological dialogue that Bacon, Descartes, and others proposed the scientific method, a method that aimed to simplify the various schools of memory, while preserving the core belief in a systematic approach to knowledge and secular outlook.
- Without the broad nature of developments in methodological thinking, without the awe inspiring feats one could perform with the art of memory, and the vision of all knowledge systematized in what we might call a computational framework, whether in Camillo’s memory theater, or in Leibniz’s universal calculus, the development of the scientific method might have had a significantly different trajectory. The art of memory provided an essential context and a vision of what would be possible with a systematic approach to knowledge. The scientific method is effectively a distillation of the principles that were largely developed and made widely known in the context of the art of memory.

This article then is organized as follows. I first give an overview of the several major periods of the *ars memorativa*, or art of memory, starting from the Classical period in ancient Greece and Rome, and then progressing through the Middle Ages and the Renaissance. The schools of thought that were in place by the end of the 16th century were the primary conceptual battlegrounds from which the scientific method emerged. I describe the major competing schools of thought of this time and provide several modern analogies that may help to shed some light into what the elite thinkers of this era were trying to accomplish. I explain Petrus Ramus’ notion of “method” and how a perhaps more proper characterization of the 17<sup>th</sup> century would be “methodological revolution” rather than “scientific revolution.” I then return to the question “what is the scientific method?” and describe how Bacon’s writings on the scientific method were one of the later arrivals onto the scene of methodological discourse and that prior interest in the various schools of thought on memory and method fundamentally enabled widespread adoption of the Baconian perspective. I will argue that what we have come to know as the scientific method is simply a small remnant of a profound set of ideas concerning the acquisition and development of knowledge, most of which have been forgotten<sup>1</sup>. Finally, I discuss Leibniz’s project for the universal calculus and examine the acceleration of mathematical thought in light of 17<sup>th</sup> century philosophy. So without further ado, let us turn our attention to ancient Greece, where as legend has it, the art of memory was born.

## THE CLASSICAL ART OF MEMORY

An apocryphal story attributes the origin of the art of memory to the lyric poet Simonides. While attending a banquet where he was to recite a poem composed in the honor of a patron, Simonides is said to have been called outside by the request of two visitors. While attending to the visitors, the banquet hall collapsed, killing everyone. Unfortunately, the bodies and faces were so mutilated that the remains could not be identified, which would preclude family members from performing the proper funeral rites. However, Simonides realized that the image of the banquet hall and those in attendance was well preserved in his mind, and by walking around the table in his mind's eye, he was able to recall the name of each person at the table and where they sat in relation to the others. Thus, he was able to identify all of the remains so that the appropriate family members could be contacted.

From this experience, legend has it, Simonides concluded that the act of memorization could be enhanced by encoding objects and concepts into spatially organized visual images. As Cicero would later write of Simonides in recounting the origin of the art of memory,

He inferred that persons desiring to train this faculty must select places and form mental images of the things they wish to remember and store those images in the places, so that the order of the places will preserve the order of the things, and the images of the things will denote the things themselves, and we shall employ the places and images respectively as a wax writing-tablet and the letters written on it.

He continues,

It has been sagaciously discerned by Simonides or else discovered by some other person, that the most complete pictures are formed in our minds of the things that have been conveyed to them and imprinted on them by the senses, but that the keenest of all our senses is the sense of sight, and that consequently perceptions received by the ears or by reflexion can be most easily retained if they are also conveyed to our minds by the mediation of the eyes. [7]

Where was this art form used? The primary domain of application was in rhetoric. Indeed, the major sources we have about the ancient art of memory come not from ancient Greece, but rather, from ancient Rome, where the method of places and images was carried on as part of the fundamental training of students in rhetoric<sup>2</sup>. The three Latin sources that largely inform our understanding of the ancient practice of this art are the *ad Herrenium*, a Roman rhetoric textbook written by an anonymous teacher written around 86 BC, Cicero's *de Oratore*, and Quintillian's *Institutio oratorio*.

Practically, as the tradition was developed, the places that people would use were often actual locations or important buildings and other architectural creations – one wonders how many facts have been encoded into mental images of the Parthenon, or later in ancient Rome, into mental images of the Colosseum. Indeed, in our modern language, we see remnants of this fascinating cognitive use of architecture. The phrase “in the first place” came into common usage when it was understood that the person speaking was using the method of places and images to encode the contents of a speech. Consequently, the audience would have understood that the “first place” and the “second place” and so on, represented actual physical locations in the mind's eye of the speaker [4, 6].

Before moving on to later centuries and the evolution of the art of memory, let me mention a critical distinction made in the rhetorical tradition, primarily to illustrate the immense capacity for memorization via symbolic representation that these ancient orators developed. The *ad Herrenium* and other texts distinguish between the “memory for things” and “memory

for words”. The “memory for things” is the classic art of memory I have described, whereby an idea, or concept of some kind is represented using a striking image. In then delivering the speech, although the order of topics is memorized, the precise words are delivered extemporaneously and not memorized a priori. On the other hand, a method that was also prescribed as a technical practice, and not for practical purpose, was the “memory for words” whereby a speech would be composed in its entirety and then memorized using the method of places and images *word for word*. This should truly astound a modern intellectual unfamiliar with these methods. The fact that the memory for words was practiced at all, even purely for the purpose of strengthening the raw mental muscles of the natural memory, demonstrates what a phenomenal capacity these ancient rhetoricians developed, not only for memory, but in developing their inner sight and ability to visualize minute details with extraordinary clarity.

## THE ART OF MEMORY DURING THE MIDDLE AGES AND THE RENAISSANCE

We see familiar prescriptions in a passage from the writings of Albertus Magnus, a 13<sup>th</sup> century Dominican Friar,

Those wishing to reminisce (i.e. wishing to do something more spiritual and intellectual than merely to remember) withdraw from the public light into obscure privacy: because in the public light the images of sensible things are scattered and their movement is confused. In obscurity, however, they are unified and are moved in order. This is why Tullius in the *ars memorandi* which he gives in the Second Rhetoric prescribes that we should imagine and seek out dark places having little light. And because reminiscence requires many images, not one, he prescribes that we should figure to ourselves through many similitudes, and unite in figures, that which we wish to retain and remember. For example, if we should wish to record what is brought against us in a law-suit, we should imagine some ram, with huge horns and testicles, coming towards us in the darkness. The horns will bring to memory our adversaries, and the testicles the dispositions of the witness. [8]

The same places and images of antiquity! It is significant that during the Middle Ages, it was mistakenly thought that the *ad Herrenium* was written by Cicero, or as he was often referred to, Tullius. As I discussed above, Cicero’s *de Oratore* is one of the few artifacts we have of the art of memory during antiquity, but the *ad Herrenium* was in fact written by an anonymous rhetoric teacher and not Tullius himself. Still, during the Middle Ages, both works were taken together as the “First and Second Rhetorics of Tullius,” a mistaken attribution which would ultimately come to lend substantial momentum to a fundamental Medieval transformation. The misattribution of the *ad Herrenium* to Cicero proved to be critical, because in *de Oratore*, he gives a great deal of attention to ethics and virtue as topical fodder for speeches. And then in the “Second Rhetoric” he gave techniques for how those topics would be properly memorized. This connection would ultimately lead to the Medieval shift of the art of memory from being a rhetorical technique to an ethical one. In particular, it is Cicero’s inclusion of memory as an integral part of the cardinal virtue of Prudence which proved vital to this evolutionary trajectory of the art of memory.

We can see this transition latent in St. Thomas Aquinas’ four precepts for memory:

Tullius (and another authority) says in his Rhetoric that memory is not only perfected from nature but also has much of art and industry; and there are four (points) through which man may profit for remembering well.

The first of these is that he should assume some convenient similitude of things which he wishes to remember; these should not be too familiar, because

we wonder more at unfamiliar things and the soul is more strongly and vehemently held to them; whence it is that we remember better things seen in childhood. It is necessary in this way to invent similitudes and images because simple and spiritual intentions slip easily from the soul unless they are as if were linked to some corporeal similitudes, because human cognition is stronger in regard to the sensibilia. Whence the memorative (power) is placed in the sensitive part of the soul.

Secondly it is necessary that a man should place in a considered order those (things) which he wishes to remember, so that from one remembered (point) progress can easily be made to the next. Whence the Philosopher<sup>3</sup> says in the book *De Memoria* ‘some men can be seen to remember from places. The cause of which is that they pass rapidly from one (step) to the next.’

Thirdly, it is necessary that a man should dwell with solicitude on, and cleave with affection to, the things which he wishes to remember; because what is strongly impressed on the soul slips less easily away from it. Whence Tullius says in his Rhetoric that ‘solicitude conserves complete figures of the simulacra’.

‘Fourthly, it is necessary that we should meditate frequently on what we wish to remember. Whence the Philosopher says in the book *De Memoria* that ‘meditation preserves memory’ because, as he says ‘custom is like nature. Thence those things which we often think about we easily remember, proceeding from one to another as though in a natural order’ [9].

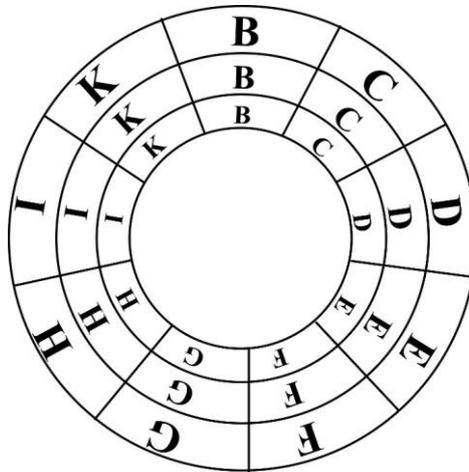
We can take away a few points from Aquinas’ precepts. First of all, we recognize, of course, as in the excerpt from Magnus, the familiar rules for places and images. And yet, whereas Magnus gives a classical example (from the *ad Herrenium* in fact) with the ram and its horns as a mnemonic image for recalling the witness in a particular set of legal proceedings, we see in Aquinas’ writing the impact of a medieval piety and the notion that “spiritual intentions,” rather than facts, can be remembered and strengthened with this practice<sup>4</sup>.

There is much more to be said about Magnus and Aquinas, but having made the basic point that the medieval practice of the mnemonic arts took on a more spiritual rather than functional character, let me now to turn to a completely separate strain of thought during the Middle Ages which would come to feature quite significantly during the 17<sup>th</sup> century. This line of thinking originated with Ramon Lull, a 13<sup>th</sup> century Franciscan friar, who developed an art of memory that proceeded along very different lines from the method of places and images, and which nevertheless had a great deal in common with the Thomist objective of strengthening one’s spiritual capacity.

Ramon Lull’s technique was distinctly evangelical in its purpose. The hope was that by arranging the fundamental concepts of Christianity on combinatory wheels, ultimately, one would be able to reduce the Christian corpus to its logical essence, and thus convince non-believers of the fundamental truth of the Christian gospel.

Lull’s combinatory wheels consisted of concentric circles, each of which was populated by symbols from the standard alphabet (see Figure 1). Each symbol was used to represent a particular concept, for example, one set of nine concepts used by Lull consisted of “Goodness, Greatness, Eternity, Power, Wisdom, Will, Virtue, Truth, Glory.” Lull then created, or imagined, concentric wheels with these concepts on each wheel, as denoted by specific symbols or characters. The wheels move independently, so by turning each wheel, in the mind’s eye of course, we arrive at a distinct set of concepts along a given column. For example, if we have a setup with 3 concentric wheels, we might imagine that in one particular

arrangement, a given column represents the concepts, “Goodness, Will, and Eternity” and by turning the second wheel by one notch we arrive at “Goodness, Virtue, Eternity.” It appears as though Lull’s combinatory wheels served two purposes. They provided a framework in which one was to structure core Christian values – presumably, some amount of effort went into the specific choice of concepts and symbols – and second, they provided a practical technique for a practitioner to first enumerate and then focus on a particular grouping of concepts. What Lull’s intent may have been in believing that this process would ultimately allow one to reduce Christian thought to its logical essentials and to free it of all logical contradictions, was that it provided a systematic method for contemplating all possible groupings of different concepts from their respective categories. Then over the course of extended contemplation and discussion, one would engage in a long-term process of re-evaluating and evolving the scriptural foundations. Lull’s combinatory wheels provided a systematic set of procedures to go about this ambitious process. Along the way, it ensured that each combination would receive careful attention, and hopefully, would draw attention to inherent logical conflicts at the basis of Christian thought.



**Figure 1.** Lullian combinatory wheel, adapted from [4; p.83].

What is remarkable about Lull’s revolving wheels, is how lacking they are in any kind of aid to memory in any spatial or visual sense<sup>5</sup>. Whereas the classical tradition seems to provide its practitioner with a powerful tool to boost the memory, Ramon Lull’s method seems to demand, from the outset, an inner sight of penetrating clarity. As a simple exercise, I encourage the reader to visualize nine letters arranged in a circle, and then try to imagine rotating the positions of the letters one by one. It is not easy! Now imagine that these letters are to be equally spaced on concentric circles, with each letter representing a concept that must be remembered on its own. Only after holding all this information in one’s mind is the practitioner in a position to begin to rotate the position of one of the wheels! If Ramon Lull alone had the ability to perform this feat, it would be noteworthy. If, as the historical record seems to indicate, there was an entire school of thought devoted to his practice, it is truly a testament to the neural plasticity of the adult brain, and its capacity to train new and unusual cognitive abilities with enormous amounts practice and concentration. Unlike the classical art, the method of Ramon Lull does not seem to take advantage of an innate capacity to remember places and images – rather it places a huge burden on our raw memory and visual ability and aims to develop them both in a single-pointed fury of devotion.

Yet another important transition took place in the philosophy and practice of the art of memory during the Renaissance. Whereas mnemonics had previously been conceived as tools for remembering, and in some cases, for instilling and intensifying religious devotion and developing certain virtues, a new notion emerged during the Renaissance. In this new school

of thought, it was put forth that relationships between the objects of mnemonic devices could be encoded into the relationships of the mnemonics themselves, and thereby give rise to the ability to understand the natural world by the strategic use of mnemonics<sup>6</sup>.

The most significant figure in this tradition was Giulio Camillo, a 16<sup>th</sup> century philosopher and scholar, who devoted the majority of his life to a project so fantastic and magnificent in scope that the entirety of Europe was abuzz in excitement of the consequences. The project was to build a theater sized repository of mnemonic figurines, physical statues which corresponded to the classical method of places and images, but which would have been standardized through great effort, and which would represent the near totality of contemporary thought. Much of Camillo's theater would have been devoted to religious and astrological concepts. The theater, which was never built, consisted of seven grades, or steps (representing the seven planets), each of which was devoted to a particular branch of knowledge. The entrant, who would have been well versed and practiced in the classical art of memory, would then be confronted with cabinets containing mnemonic images representing some specific concept, fact, or idea. In addition, there were numerous drawers which were filled with papers of speeches from Cicero relating to the topical matters of the images depicted in the surrounding areas. Viglius Zuichemus, a Dutch statesman, describes the theater and his meeting with Camillo in a letter to his contemporary, Erasmus:

The work is of wood, marked with many images, and full of little boxes; there are various orders and grades in it. He gives a place to each individual figure and ornament, and he showed me such a mass of papers that, though I always heard that Cicero was the fountain of richest eloquence, scarcely would I have thought that one author could contain so much or that so many volumes could be pieced together out of his writings. I wrote to you before the name of the author who is called Julius Camillus. He stammers badly and speaks Latin with difficulty, excusing himself with the pretext that through continually using his pen he has nearly lost the use of speech. He is said however to be good in the vernacular which he has taught at some time in Bologna. When I asked him concerning the meaning of the work, its plan and results – speaking religiously and as though stupefied by the miraculousness of the thing – he threw before me some papers, and recited them so that he expressed the numbers, clauses, and the artifices of the Italian style, yet slightly unevenly because of the impediment in his speech. The King is said to be urging that he should return to France with the magnificent work. But since the King wished that all the writing should be translated to French, for which he had tried an interpreter and scribe, he said that he thought he would defer his journey rather than exhibit an imperfect work. He calls this theater of his by many names, saying now that it is a built or constructed mind and soul, and now that it is a windowed one. He pretends that all things that the human mind can conceive and which we cannot see with the corporeal eye, after being collected together by diligent meditation may be expressed by certain corporeal signs in such a way that the beholder may at once perceive with his eyes everything that is otherwise hidden in the depths of the human mind. And it is because of this corporeal looking that he calls it a theater [10].

Camillo's theater represents a third and distinct phase of the art of memory. What began in antiquity as a functional practice for memorizing the contents of speeches in the rhetorical tradition, then took on a spiritual character during the Middle Ages in the hands of St. Thomas Aquinas and the Dominican tradition. In a sense, Camillo's mnemonic theater fuses the classical art of memory with the intent of Ramon Lull's quite distinct practice of religious concepts

encoded into combinatory wheels. Camillo's theater, while still in the tradition of the method of place and images, represents a vision of knowledge systematized to a heroic degree, not solely for the purpose of memory, but for understanding the structure of knowledge itself and for accurately reflecting the basic principles of the cosmos in its mnemonic organization.

## THE 17<sup>TH</sup> CENTURY AND A CONCEPTUAL REINVENTION

Entering the 17<sup>th</sup> century, we are confronted with the following raw materials in the world of method and memorization. From ancient Greece and ancient Rome, via the rhetorical tradition and through the work of the Dominican order comes the tradition of the *ars memorativa* and the systematic and creative use of mnemonics for the purpose of memory, and in a closely related capacity, for ethical development. On the other hand, from Ramon Lull, who we recall as early as the 13<sup>th</sup> century was developing his own system of symbolic representation of religious concepts, we have the notion of an intensive practice of memory, but without the associated system of images. From Giulio Camillo, we have the Renaissance transformation of the classical art, with its rules for images and places well preserved, but now with the additional notion that mnemonic images could also be used to represent the structure of knowledge and of Nature itself.

And of course, memory through "mere" repetition is an age old notion that has no name to which we can attribute a founder, although it was more precisely formulated by a man who is critical in understanding the 17<sup>th</sup> century and the origin of the scientific method, Petrus Ramus. Ramus was a 16<sup>th</sup> century educational reformer who was particularly concerned with re-examining the ways in which subjects were to be memorized. In particular, Ramus made the dramatic, iconoclastic move of eliminating memory as part of rhetoric. With this gesture, he was able to distance himself from the mnemonic practice of places and images, which he replaced with a new method, called the "dialectic method." In this method, a subject is first structured in a logical manner by proceeding from the most general concepts to the most specific. The content is then memorized by the standard practice of repetition.

Part of Ramus' motivation as a reformer related to a specific set of religious objections to the classical art of memory which stem from the Old Testament:

Take ye therefore good heed unto yourselves; for ye saw no manner of similitude on the day that the Lord spoke unto you in Horeb out of the midst of the fire: Lest ye corrupt yourselves, and make you a graven image, the similitude of any figure the likeness of male or female . . . And lest thou lift up thine eyes, unto heaven, and when thou seest the sun, and the moon, and the stars, even all the host of heaven, shouldst be driven to worship them . . . [11].

According to Yates, this prohibition of graven images, taken from the fourth chapter of Deuteronomy, was interpreted by Ramus as applying to the classical art of memory as well. The rhetorical tradition quite actively advocated the use of lewd and grotesque images by which to excite the imagination and empower the memory, which to Ramus, was tantamount to a systematic technique for polluting one's mind.

It is not too difficult to sympathize with Ramus when considering some of the writings of his predecessors. Pietro de Ravenna, a 15<sup>th</sup> century jurist, ardent self-promoter, and evangelist of the method of places and images wrote the following as a suggested set of techniques for inventing more effective mnemonics:

I usually fill my memory-palaces with the images of beautiful women, which excite my memory . . . and believe me: when I use beautiful women as memory images, I find it much easier to arrange and repeat the notions which I have entrusted to those places. You now have a most useful secret of artificial

memory, a secret which I have (through modesty) long remained silent about: if you wish to remember quickly, dispose the images of the most beautiful virgins into memory places; the memory is marvelously excited by images of women. . . This precept is useless to those who dislike women and they will find it very difficult to gather the fruits of this art [12].

It is worth noting that there is a purely secular counterpart to the Ramist objection, not on the grounds of spiritual pollution, but rather, conceptual interference. As I have described, and as an any reader who chooses to experiment with these techniques will see for him or herself, the classical method of places and images gains its strength by taking advantage of vivid conceptual associations, which almost always have very little in common with the specific facts or ideas that are being remembered. It is necessarily the case then, that by using these techniques extensively, one is creating a vast array of associations that are completely arbitrary and have no resemblance to the logical structure of the content being memorized.

In attempting to circumvent these stray associations, whether on religious or secular grounds, Ramus introduced a concept that will prove to be critical in understanding the 17<sup>th</sup> century and the scientific revolution. Ramus was the first thinker to popularize the word “method” [4, p. 369]. In the modern world, we understand the connotation of “method” as referring to an orderly, procedural practice, but in the 16<sup>th</sup> century when Ramus began to popularize this word, recall that he was advocating a specific type of method – the dialectic method – which stood in contrast to the “mnemonic method,” i.e., the classical art of memory, and the “method of Ramon Lull,” consisting of combinatory wheels with associated symbols.

To understand the specific connotation of this word as it would have been understood by Ramus and his contemporaries, I propose that we examine the modern word *algorithm*. The Oxford-English Dictionary defines an algorithm to be,

A procedure or set of rules used in calculation and in problem solving; a precisely defined set of mathematical or logical operators for the performance of a particular task [13].

I think many would agree that something is lacking in this definition. Specifically, the word algorithm is a fairly recent word, and yet this definition describes a notion that has existed for thousands of years. In particular, Euclid’s algorithm for finding the greatest common divisor of two numbers dates back as far as 300 BC. What seems to be critically missing from the definition, is the additional *connotation* the word algorithm gains when used in reference to modern computing technology. Certainly, young children can learn algorithms, say for performing long division, well before they are exposed to computer programming, and teaching them the word algorithm in this context would add very little to their conceptual maturity. However, the word itself gains substantial depth with the additional notion that a computing device can perform a set of instructions thousands, millions, and billions of times, with a precision and accuracy that no human could otherwise accomplish. This additional connotation, which is experienced through interacting with computers, through film and television, and which is not captured in the purely dictionary definition, is critical, and we can re-examine the word “method,” as used by Ramus, in a similar light.

While Ramus may have objected to the classical use of images and places, his own dialectical method – one particular kind of “method” – certainly profited from the connotation that the art of memory carried with it, namely as a systematic procedure for the memorization of knowledge. But furthermore, Ramus’ dialectic method had also something in common with the “method of Ramon Lull,” as both methods aimed to distill a particular knowledge base, in Lull’s case the Christian doctrine, to its logical essence. Indeed, Lullian combinatory wheels may have played an essential role in lending a practical set of associations to the word “method,” in the same way that computing technology is essential in understanding the word

algorithm in modern times. Recall again, the standard Lullian setup, say for example, consisting of 3 concentric circles with 9 Greek symbols designating the concepts Goodness, Greatness, Eternity, Power, Wisdom, Will, Virtue and Glory. Lull's method for enumerating all  $(\quad) = 84$  distinct combinations provides a systematic procedure for ensuring that in the process of contemplating the implications and interrelation of each concept with each other, no stone would be left unturned.

We have seen thus far that the notion of "method" unified the diverse strains of thought related to memory, and that in an important way, the pure act of memorization was secondary to the notion of method as a systematic procedure for acquiring knowledge and investigating Nature. Amid these turbulent conceptual battles being fought on the grounds of the mnemonic, dialectic, and Lullian methods, with disagreements about the very foundations of method, there were two great mollifiers and unifiers who played a crucial role in the emergence of the mathematical and scientific methods, Giordano Bruno and Gottfried Leibniz.

Bruno and Leibniz were both thoroughly versed in all branches of methodological thinking, from the classical method of places and images, to the method of Ramon Lull, to the dialectic method of Ramus. Born four years after the death of Giulio Camillo, Bruno was trained in a Dominican convent in Naples, where he would have certainly been exposed to the method of places and images from the *ad Herrenium*, and to the work of his predecessor, St. Thomas Aquinas. Bruno's first treatise on memory, *De umbris idearum* was published in 1582, and was dedicated to the King of France:

I gained such a name that the King Henri III summoned me one day and asked me whether the memory which I had and which I taught was a natural memory or obtained by magic art; I proved to him that it was not obtained by magic art but by science. After that I printed a book on memory entitled *De umbris idearum* which I dedicated to His Majesty, whereupon he made me an endowed reader [14].

Bruno's early work had much in common with the tradition of Camillo, with a significant emphasis on the use of places and images for astrological purposes. Having been brought up in a world in which Camillo's influence would have been strongly felt and talk of his magnificent theater quite active, and in addition, belonging to the same Dominican order as St. Thomas Aquinas, Bruno was heir to the most significant historical developments stemming from the classical art of memory. But his historical position is significantly colored by exposure to that other thread of the art of memory, the method of Ramon Lull. Indeed, Paris, which at that time was far, but certainly accessible from Bruno's hometown of Naples, was the epicenter of 16<sup>th</sup> century Lullism, and it was in combining these two traditions that Bruno would make his mark.

Bruno took the bold move of unifying both traditions by starting with Lullian combinatory wheels, but using classical mnemonic images instead of the standard alphabet for representing concepts on each locus. Thus, he replaced the architectural component of the method of places and images by a Lullian combinatory wheel. Furthermore, in contrast to Ramon Lull, Bruno's objectives were not strictly religious in nature. Indeed, he had a great deal in common with Giulio Camillo, and we can see the extent of this vision in the shocking size and complexity of Bruno's wheels. In a particular example extracted from one of his treatises, we see detailed a combinatory wheel with 30 divisions, each of which is further divided into 5 parts, giving 150 divisions in total. The lists that he includes in the book are sets of 150 images each, which are to be ordered on the wheels in a Lullian fashion [4; pp.213-223, 5; pp.87-88]. In the remainder of this article, I will refer to Bruno's system as the hybrid mnemonic-Lullian method.

The complexity of such a system is truly appalling. I invite the reader, if he or she has not already done so, to attempt the simple exercise I outlined earlier of visualizing letters

arranged in a circle. Having gotten to the point of visualizing this structure clearly, now attempt to rotate the position of each letter. If this exercise seems demanding, imagine what kind of mental facility would be required to manipulate the hundreds of detailed images populating a Brunian mnemonic wheel! As I stated of Ramon Lull, if Bruno alone could visualize an object of such stunning complexity, it would have been a noteworthy accomplishment, if only for demonstrating the human mind's capacity for training unusual cognitive abilities.

It may surprise readers that Gottfried Leibniz, who most know as being co-inventor of the infinitesimal calculus, was one of the foremost figures in methodological innovation. Indeed, his role during the 17<sup>th</sup> century paralleled that of Bruno, and whereas Bruno attempted to unify the Lullian and mnemonic traditions, Leibniz's primary effort, indeed his overarching vision, was an effort to unify the mnemonic method with the dialectic method of Petrus Ramus<sup>7</sup>. Certainly Leibniz was intimately familiar with the Lullian tradition and also with Bruno's attempts at unification. The "universal calculus" was Leibniz's primary aim and it borrowed from both the dialectical and mnemonic traditions in the following manner. First, an encyclopedia was to be constructed covering the entire domain of human thought, from science, to religion, to law. In the spirit of the dialectic method, the encyclopedia would be carefully assembled so as to reflect the natural logical structure of each discipline<sup>8</sup>. Finally, drawing from the mnemonic method, symbols would be constructed from each core concept, from which, Leibniz hoped, a *universal calculus* would emerge in which logical contradictions could be eliminated from the entirety of human thought and in which all questions, whether legal, scientific, or religious could be answered by *computing the answer* via manipulation of the associated symbolic infrastructure. As we see in the following passage, it is clear that Leibniz viewed this project as requiring permanent, ongoing evolution, progressing hand in hand with the development of scientific knowledge:

Although this language (the universal calculus) depends on true philosophy, it does not depend on its perfection. Let me just say this: this language can be constructed despite the fact that philosophy is not perfect. The language will develop as scientific knowledge develops. While we are waiting, it will be a miraculous aid: to help us understand what we already know, and to describe what we do not know, and help us to find the means to obtain it, but above all it will help us to eliminate and extinguish the controversial arguments which depend on reasons, because once we have realized this language, calculating and reasoning will be the same thing [15].

Leibniz seemed to be keenly aware of the importance of proper notation<sup>9</sup>, and like Ramus, he also seemed to show concern for the potential for conceptual interference that the mnemonic method necessarily entailed. But whereas Ramus chose to abandon the notion of symbolic representation entirely, Leibniz continued to express considerable confidence that one could find a set of symbols that would both enable the memory and serve as a proper foundation for what we might call a "calculus of concepts." Indeed, Leibniz hoped that by borrowing geometric elements from the Egyptian alphabet, or pictorial elements from the Chinese alphabet, he would arrive at a set of symbols which would precisely satisfy these competing constraints [4; p.381, 5; pp.179-180].

## **THE EMERGENCE OF THE SCIENTIFIC METHOD**

We are now in a position to tackle the question I posed at the beginning of this article, "what is the scientific method?" As I stated in the introduction, the standard answer is not entirely inaccurate. As the Oxford English Dictionary states, the scientific method could certainly be

defined as “a method of procedure that has characterized natural science since the 17<sup>th</sup> century, consisting in systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses.” And Francis Bacon was certainly among the first to articulate these principles [16].

What may surprise readers is that Bacon was not only aware of the art of memory, but also quite well versed in all aspects of methodological thinking. Indeed, in *De augmentis scientiarum*, we see a very familiar description of the method of place and images:

Emblems bring down intellectual to sensible things; for what is sensible always strikes the memory stronger, and sooner impresses itself than the intellectual . . . And therefore it is easier to retain the image of a sportsman hunting the hare, of an apothecary ranging his boxes, an orator making a speech, a boy repeating verses, or a player acting his parts, than the corresponding notions of invention, disposition, elocution, memory, action [17].

But as a historical figure, Bacon closely resembled Petrus Ramus as an educational theorist, and in *The Advancement of Learning*, the art of memory is featured as one of the arts and sciences in need of reform. In particular, Bacon’s primary objection was that the art of memory could be used for unnecessarily acquiring massive amounts of information simply for the purposes of impressing people, and that this kind of effort would be more usefully directed at advancing intellectual projects.

Renes Descartes, another central figure in the emergence of the scientific method, also took on a Baconian interest in reforming the art of memory. Perhaps coming more prominently from the tradition of natural philosophy and mathematics, Descartes was interested in using the art of memory for understanding causality. In this fascinating passage, we see Descartes drawing a fundamental connection between the mnemonic method and reductionism:

On reading through Schenkel’s<sup>10</sup> profitable trifles (in the book *De arte memoria*) I thought of an easy way of making myself master of all that I discovered through the imagination. This would be done through the reduction of things to their causes. Since all can be reduced to one it is obviously not necessary to remember all the sciences. When one understands the causes all vanished images can be easily found again in the brain through the impression of the cause. This is the true art of memory and it is plain contrary to (Schenkel’s) nebulous notions. Not that his (art) is without effect, but it occupies the whole space with too many things and not in the right order. The right order is that the images should be formed in dependence on one another. He (Schenkel) omits this which is the key to the whole mystery.

I have thought of another way; that out of unconnected images should be composed new images common to them all, or that one image should be made which should have reference not only to the one nearest to it but to them all – so that the fifth should refer to the first through a spear thrown to the ground, the middle one through a ladder on which they descend, the second one through an arrow thrown at it, and similarly the third should be connected in some way real or fictitious [18].

As has been the common theme throughout this article, we see that the art of memory, sagaciously discerned by Simonides, or else discovered by some other person, has once again captivated the imagination of another great figure of the Age of Reason. With Bacon and Descartes as my final historical examples, I hope to have convinced the reader at the very least, that the classical art of memory and its many descendants played a critical role in

Western intellectual history and in the lives of many prominent thinkers<sup>12</sup>. Let us now return to the origin purpose of this article and re-examine the question “what is the scientific method?”

Most scientists will likely agree with the statement that good scientific practice largely involves perseverance, determination, curiosity, systematic thinking, patience, and experience. And I suspect that most will agree that many of these qualities are the critical determinants of success in other, non-scientific fields as well. Conversely, bad science can be characterized by efforts lacking in systematization, lacking in curiosity, impatience, and a lack of experience. And I think most would agree that these qualities are learned, as I stated early on in this article, by simply participating in the frontiers of research under the guidance of more experienced researchers.

The observation that scientific maturity is largely communicated via implicit, cultural knowledge suggests that the development of the scientific method was not a discrete event. As I stated in the introduction, we can certainly point to numerous examples of principled scientific reasoning well before the 16<sup>th</sup> and 17<sup>th</sup> centuries. Indeed it would be difficult to imagine that hypothesis-driven investigation in some limited form or another was not always part of human society. To properly re-contextualize the original question, therefore, it seems more appropriate to ask “how and why did the adoption of the scientific method accelerate during the 16<sup>th</sup> and 17<sup>th</sup> centuries?”

There is no doubt that the more commonly attributed elements of the scientific revolution were significant – a vision of the natural sciences built upon hypothesis-driven investigation, the establishment of a scientific journal, and the critical support of wealthy patrons, etc. But the question remains as to why these factors gained traction during this time period when they had not before. Without a critical mass of scientific accomplishments, these principles, which seem elementary and self-evident to us, would have been rather abstract to individuals of that era. What would have been the motivation to adopt a worldview which amounts to a highly philosophical, meta-scientific position towards the systematic pursuit of knowledge?

It is specifically in this context, that is, in providing momentum and an inspiring vision to motivate the widespread adoption of a scientific viewpoint that I believe the art of memory to have played an instrumental role. In particular, the cultural and institutional circumstances of 17<sup>th</sup> century Europe provide an answer to the question of how philosophical argumentation concerning the principles of reason and using rational thought to understand Nature could have taken root in a world without the academic infrastructure that we have today. There is a compelling argument to be made that a critical enabler of scientific thought was that the two medieval schools of memory which would figure prominently in the 17<sup>th</sup> century, the classical art as preserved by St. Thomas Aquinas and the Dominican order, and the method of Ramon Lull preserved by the Franciscans, would have been well known throughout Europe, given the geographic flexibility of the friars. Thus, not only would the most recent developments be intelligible to newcomers, for instance Bruno’s hybrid mnemonic-Lullian method, or Leibniz’s hybrid mnemonic-dialectic method, but in addition, there was a specific set of techniques that one could put into practice in order to understand the implications of the larger philosophical principles and world views being advocated.

Without a concrete set of actions for listeners to engage with, it seems highly unlikely that the Baconian perspective would have resonated with a sufficiently large group of people to precipitate institutional or cultural change. Imagine being a moderately or even highly educated member of 17th century European society. You attend a lecture by Bacon, or one of his followers, on the virtues of reason and the notion that humanity is best served by the most educated and talented minds directing their efforts toward the natural sciences. Even if you were inspired by these words, what is the next step? What does one *do* next? In a world in

which survival itself was a more serious undertaking, in a continent torn by political instability, and in which recurring plagues decimated major cities, it seems quite likely that a person's attention would be strongly pulled back to practical realities.

On the other hand not only did the mnemonic method, the dialectic method, and the method of Ramon Lull all carry with them analogous rhetorical potency, they slowly began to be directed towards the study of the natural world and the development of the arts and sciences. In addition, all of these methods were concrete practices that would have been universally understood, and which provided an inspiring vision of what could be accomplished with a systematic, rational approach to the pursuit of knowledge. If one was familiar with the mnemonic method, and was then exposed to Leibniz's notion of the universal calculus, one could, in principle, begin immediately experimenting with these new ideas, or at the very least contemplate them and discuss them with others. One could fall asleep at night, slowly pacing through thousands of memory loci, imagining that rather than being used for memorizing the words of Cicero, these same loci could be mathematical symbols which encoded a calculus of the natural sciences. These *internal experiences* would have been powerful, and thus, I believe that a significant agent of philosophical change in the 17<sup>th</sup> century was the general notion of "method," and in particular, the descendants of the classical art of memory.

If these powerful mental abilities were directed solely at rhetoric, we can imagine that they would have no bearing on future scientific institutions. But the historical record indicates otherwise. Indeed, over the course of nearly two millennia, we see that the art of memory went through multiple transformations and became increasingly directed at general knowledge and the unearthing of natural principles. And in the eyes of some of the greatest minds of the Western intellectual tradition, the art of memory was viewed as something substantially more general and versatile than merely a tool for memorization. Indeed, nearly three centuries after the invention of the printing press, we see that the art of memory continued to be investigated, developed and improved upon, in the diverse incarnations of the mnemonic method, the dialectic method, the method of Ramon Lull, the mnemonic-Lullian method of Giordano Bruno, and the mnemonic-dialectic method of Gottfried Leibniz. And not only were those pioneers traditionally associated with the rise of the scientific method, Francis Bacon and Renes Descartes, well-versed practitioners of these "other" kinds of method, but their outlook and vision for the continued development of intellectual institutions were formulated in reference to the memory arts.

One way to conceptualize the intellectual milieu of this time period is to recognize that the phrase "scientific method" is in fact a compound construction consisting of the two words "scientific" and "method." Just as "organic chemistry" is a specific kind of chemistry as distinguished from "physical chemistry" and just as "quantum mechanics" is a specific kind of mechanics as distinguished from "classical mechanics," the "scientific method" is a specific kind of method which at one point in time was distinct from other kinds of method such as the dialectic method, the method of Ramon Lull, the mnemonic method, and so on. As I have argued, scientific thinking has always been part of human society – what took place during the 17<sup>th</sup> century was that scientific thinking became fused with methodological thinking. In other words, it was the art of memory that inspired the "method" in "scientific method."

In a sense, the principles articulated by Bacon, Descartes, and others, were something of the aftermath of what should perhaps be called the "methodological revolution" rather than the "scientific revolution." The Baconian school was in a position to survey the massive intellectual transformation that was created by the diverse manifestations of the memory arts, and these writings serve to distill the basic qualities embodied by the widespread practices of "method," and to focus these efforts even more strongly in the direction of the natural

sciences. But without the prior history of the mnemonic method and its many descendants, this perspective may not have had nearly the same impact.

## **THE ART OF MEMORY AND THE GROWTH OF THE MATHEMATICAL METHOD**

I hope to have convinced the reader of the significant impact the art of memory has had on Western intellectual history<sup>12</sup>. It is a practice that has been the foundation of so many institutions and cultural movements, from Greek and Roman rhetorical students and statesman in antiquity, to Franciscan and Dominican friars in the Middle Ages, to the natural philosophers and scientific trailblazers of the 17<sup>th</sup> century. For an idea that has so profoundly sustained and nurtured Western civilization, it made a truly graceful exit.

But perhaps it is an opportune moment in history to reconsider its value, and in the remainder of this article, I will examine the role that the art of memory may have played in the growth of mathematical thought. In particular, I will examine Gottfried Leibniz's project to construct a *universal calculus* which would resolve logical conflict in all areas of knowledge. My primary aim in this section is to try to understand details of Leibniz's research agenda in greater depth, and in particular, to attempt to articulate what to modern mathematical scientists appears to be a rather peculiar set of beliefs about mnemonics and their potential for conceptual abstraction. Along the way – in a sense, attempting to view the world as Leibniz or one of his followers would – I will argue the following:

- Mnemonics and mathematics both share the core property of being symbolic representations of concepts.
- In an era where mathematical thought was substantially less developed, and in particular, where there were far fewer applications of mathematics to the natural sciences, the difference between mnemonics and mathematics may have been viewed as being rather small.
- Leibniz, in particular, may have viewed his work on the infinitesimal calculus as simply a “toy” problem in the much larger vision of the universal calculus, in which mnemonics would have played a critical role in creating a symbolic representation of all human knowledge. That is, mnemonics were viewed as being a potential basis for physical theories, along side more traditional mathematics.
- Conversely, the motivation to pursue mathematical theories for physical phenomena may have gained additional momentum via its affiliation with mnemonics, and in particular, the mnemonic theater of Giulio Camillo, Bruno's hybrid mnemonic-Lullian method, and Leibniz's hybrid mnemonic-dialectic method. In other words, even though the mnemonic method ultimately did not give rise to physical models, there was a widespread vision and tremendous confidence that it would someday. This vision may have helped to accelerate interest in pursuing mathematical approaches to model physical phenomena.
- The 17<sup>th</sup> century vision of mnemonics as being closely related to mathematics may provide us with a novel attack on the foundational questions of “what is mathematics?” or “why can mathematics be used to model natural phenomena?”

We live in an era where mathematics and the physical sciences have advanced to such an extent that we are able to make conceptual distinctions that would seem quite foreign to even the most sophisticated minds of a few centuries ago. Mathematical researchers categorize analysis, algebra, and topology as clearly distinct topics from which interdisciplinary work in areas such as algebraic topology or analytic number theory can arise. Professional scientists can be heard remarking about the “completely different planets” occupied by theoretical physics and mathematical physics – distinctions that are often difficult for bright undergraduates to appreciate, let alone a scientist from several centuries ago when even the most primitive

concepts underlying modern physics were in their nascent stages of formalization.

In the modern era, we have seen so many examples of mathematical success, both in the rich conceptual structures of pure mathematics, and in the application of mathematics to the physical sciences, that we can lose sight of the fact that there may have been a time when one might have been justified in maintaining some amount of doubt about the long term viability of mathematics itself as a worthy topic of investigation. The world today has been so fundamentally imbued with the successes of mathematics, from the earth shattering precision of quantum electrodynamics and general relativity, to the society transforming potential of the Internet, medical imaging, and autonomous vehicles, that it is easy to forget to ask the question of why mathematics *works at all* in describing natural phenomena. And in an earlier era with far fewer examples of mathematical success in the natural sciences, what might have been the competing schools of thought that scientific minds might have otherwise invested their efforts in?

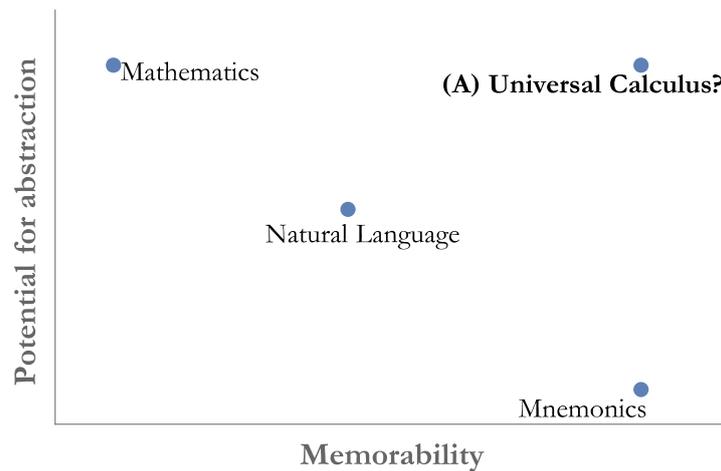
In the article “On the unreasonable effectiveness of mathematics in the natural sciences,” physicist and quantum mechanics pioneer, Eugene Wigner, writes the following:

Most of what will be said on these questions will not be new; it has probably occurred to most scientists in one form or another. My principal aim is to illuminate it from several sides. The first point is that the enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious and that there is no rational explanation for it. Second, it is just this uncanny usefulness of mathematical concepts that raises the question of the uniqueness of our physical theories. In order to establish the first point, that mathematics plays an unreasonably important role in physics, it will be useful to say a few words on the question, “What is mathematics?”, then, “What is physics?”, then, how mathematics enters physical theories, and last, why the success of mathematics in its role in physics appears so baffling [22].

One interesting hypothesis that Wigner suggests is that the success of mathematics in contemporary science is partly due to a selective bias. That is, as scientists, we have chosen, somewhat by necessity, to examine a subset of physical phenomena which are amenable to mathematical description. In understanding how mathematical thought, and particularly its role in the physical sciences, could have taken root, it is worth examining the most general property of mathematics, namely as a symbolic alphabet for representing concepts. What I will take as a starting point for the remainder of this discussion is the observation that the art of memory should also be thought of as a symbolic alphabet for representing concepts. While this may not have been the perspective taken during antiquity, it is in the Renaissance transformation of the art of memory and the memory theater of Giulio Camillo that we see the worldview develop that mnemonic images could be used to represent the very structure of reality, in other words, *as a potential basis for physical theories*. Consequently, I conjecture that in an era where the notion of mathematical modeling and symbolic representation were in their infancy, the distinction between the memory palaces of the mnemonic method and mathematics itself may have been viewed as quite small, or rather bearing the kind of resemblance that a modern mathematician might ascribe to analysis and algebra, or between a mathematical derivation in a physics paper and a computer simulation of the same system. That is, these two forms of symbolic systems had enough in common that contrasting them and advocating the use of one over the other would have been a reasonable thing to do. For things substantially more dissimilar, it would be completely unnecessary – one does not need to advocate the use of a broom instead of an apple for sweeping the floor, but might yell at someone about to use a mop.

I raise this point because I believe it provides us with a novel approach to attacking the

question “what is mathematics?” or “why can mathematics be used to describe physical phenomena?” Suppose we would like to compare symbolic systems by the degree to which they possess two different qualities – the capacity for aiding human memory and the capacity for building up successive layers of conceptual abstraction. On such a scale, we would imagine that the symbols of the mnemonic method have an extremely high degree of memorability, but low capacity for abstraction, whereas the symbols of usual mathematics have a high degree of inherent potential for conceptual abstraction, but a moderately low inherent tendency toward precise recall. It seems that human natural language lies somewhere in between the mnemonic method and mathematics on both scales (see Figure 2).



**Figure 2.** Cartoon diagram showing where different symbolic systems lie with respect to the criteria of memorability and potential for conceptual abstraction. The point (A) in the upper-right indicates a hypothetical system combining the high degree of memorability of mnemonics with the abstract potential of mathematics – precisely what Leibniz was attempting to achieve with the universal calculus.

We might then ask what other points on this graph could represent. In particular, is it possible to develop a symbolic system which has the inherent capacity to aid the human memory, as the mnemonic method does, but which, like mathematics, also has the potential for building many successive layers of abstraction? Indeed, I believe that this is exactly the question that Leibniz was tackling in formulating the notion of the universal calculus. Recall that the agenda of the universal calculus was to first build an encyclopedia, in a manner akin to the dialectic method, and to then assign a symbol for each concept. But Leibniz was keenly aware of the controversy surrounding the mnemonic method’s problematic side effects for generating conceptual interference, and was looking to the Chinese and Egyptian alphabets as possible sources of inspiration for how he could retain the memory aiding features of the mnemonic method, while eliminating the unnecessary conceptual interference inherent in complicated and fantastic images.

I find Leibniz’s vision to be truly romantic and beautiful, and yet, I also have an inherent skepticism that it would have worked. Suppose, for example, that in the dialectic tradition, we create a structured list of all biological knowledge relevant to frogs. As Ramus advises us, we organize the information starting with the most general and proceeding towards the most specific. Now, following Leibniz, we attempt to create a symbolic representation for each of these concepts (recall the example given earlier of associating predicates with prime numbers). Would we ultimately arrive at a “calculus of frogs?” This is difficult to imagine, and yet it seems that something along these lines is what Leibniz was trying to accomplish. It is possible that enough has been forgotten of the thought process of the era that what I have written here is a misinterpretation of Leibniz’s agenda. In that case, it seems like ripe fodder

for further historical work to re-examine the primary sources to understand why Leibniz, a man who had just created one of the most powerful mathematical structures the world has seen, had tremendous optimism that the universal calculus based on the hybrid mnemonic-dialectic method would have given rise to similar levels of conceptual abstraction and capacity for prediction as the infinitesimal calculus. We see in the following inspired words, the zeal with which Leibniz pursued this vision:

My invention contains all the functions of reason: it is a judge for controversies; an interpreter of notions; a scale for weighing probabilities; a compass which guides us through the ocean of experience; an inventory of things; a table of thoughts; a microscope for scrutinizing things close at hand; an innocent magic; a non-chimerical cabala; a writing which everyone can read in his own language; and finally a language which can be learnt in a few weeks, traveling swiftly across the world, carrying the true religion with it, wherever it goes [23].

Assuming for the moment that Leibniz's hybrid mnemonic-dialectic method is unlikely to have given rise to a universal calculus in the way his more mathematical work gave rise to the integral and differential calculus, let me now state a question that I believe to be an alternative approach to attacking the questions "what is mathematics?" or "why can mathematics be used to describe natural phenomena?" The question is the following: why does mathematics keep going where Leibniz's hybrid mnemonic-dialectic method stops? That is, it appears as though both mathematics and mnemonics share the core property of being symbolic representations of concepts, however, the latent potential in mathematics for building successive layers of abstraction is infinite, whereas with mnemonics, even when used in conjunction with the logical structure of a dialectically organized encyclopedia, it seems as though it doesn't quite get off the ground. And yet for some of the greatest minds of the scientific revolution – Gottfried Leibniz in particular, but also Rene Descartes – there was a tremendous amount of confidence that mnemonics would give rise to precisely the kind of conceptual power that mathematics has ultimately given us, and for which there was only a small amount of evidence at the time.

There is one obvious hypothesis for why mnemonics lack the abstract potential of mathematics, which I think ultimately fails, but is worth examining anyway. As I have stated throughout this article, one problematic feature of mnemonics is the built-in conceptual interference that the techniques give rise to. In fact, it seems as though conceptual interference is precisely what enables the powerful ability to recall minutiae with such accuracy. Suppose I want to remember that the adjective "feline" means "possessing the qualities of a cat." I can create a mnemonic that uses the sound of the prefix "fe" to remind me of the element iron, and I can imagine a huge iron cat, possibly two huge iron cats standing on either side of a large iron gate which stands at the entrance to a large palace, and in which many cats are standing in "line" to visit. Using this association, I can remember that "feline" means "possessing the qualities of a cat." However, for the most part, the mnemonic itself consists of information largely irrelevant to anything related to cats, and so in a fundamental way, this image is incredibly misleading, although our brains seem to possess the remarkable ability to know to ignore the vast majority of the information contained in the mnemonic image. I am unlikely to begin to think that cats are able to stand in line or that there are palaces devoted to cats by virtue of using this particular mnemonic.

So it seems as though mnemonics gain their power by taking advantage of obscure and fantastic connections between word and object and that while this may aid the memory, it may also be what inhibits further conceptual potential. In fact, as I have described earlier in

the article, this objection is precisely what was raised by Ramus and his disciples, the proponents of the dialectic method. Of course, their objection was not entirely an intellectual one – it was that the use of the mnemonic method would pollute the mind with unclean images and that this would corrupt one’s spiritual character. But it seems that for other intellectuals, there would have been completely secular reasons for opposing use of the mnemonic method, and conceptual interference was the basis for this objection.

Can we conclude then that the reason that mnemonics do not reach the conceptual capacity of mathematics is that the information content of mnemonics is dominated by concepts peripheral to the core notion to be remembered? This is possible, however, it is worth mentioning that standard mathematical notation itself does contain some amount of conceptual interference, albeit substantially less than mnemonics. For example, in quantum mechanics, we frequently denote the system density matrix with the symbol  $\rho$ . The symbol  $\rho$  is round – does that mean the density matrix itself is also round? Of course, not, it is merely a symbol and could just as well be represented by  $\psi$ ,  $\Phi$ , or  $\div$ . Probability distributions are often times denoted by  $\pi$ . Does that mean that they have some deep connection with the ratio of a circle’s circumference to its diameter? Again, of course, not, but it is clear from this example that mathematical notation can also give rise to conceptual interference, albeit substantially less than the elaborate scenery of the mnemonic method.

Before bringing this discussion to a close, let me restate my primary purpose in this section. For modern intellectuals, Leibniz’s agenda of constructing the universal calculus via the mnemonic-dialectic method sounds extremely strange. How could mnemonic images possibly give rise to either conceptual abstraction or predictive power when they are simply one-to-one maps from some list of facts? And yet it appears as though this viewpoint had quite a bit of support during the 17<sup>th</sup> century. In particular, it is noteworthy that Leibniz, who stands among elite company in having developed the infinitesimal calculus, had the most confidence of any thinker of the era – it would be surprising if the obvious objections to this project were lost on him. Therefore, I believe that the question “why does mathematics have a near infinite capacity for conceptual abstraction whereas mnemonics do not?” is a new one and may provide a fresh perspective on how to attack some long-standing questions in the foundations of mathematics. Perhaps in trying to articulate explicitly the many implicit assumptions in our worldview, we will discover something novel and of practical relevance to contemporary intellectual questions.

## **ACKNOWLEDGEMENTS**

I would like to thank Michael Rosen, Aaswath Raman, and Doug Bemis for insightful discussions and critical reading of the manuscript.

## **REMARKS**

<sup>1</sup>Joshua Foer’s memoir *Moonwalking with Einstein* is a brilliant and engaging first person account of the “competitive memory circuit,” where many of these ideas remain alive [6].

<sup>2</sup>I will use the phrases “method of places and images,” *ars memorativa*, and “method of loci” more or less interchangeably in his article.

<sup>3</sup>These quotations are a reference to Aristotle.

<sup>4</sup>Of course, Magnus also hints at this idea in the preceding quotation: “Those wishing to reminisce (i.e. wishing to do something more spiritual and intellectual than merely to remember) . . . ”

<sup>5</sup>Lull did develop another method which is more reminiscent of the classical art, in which different branches of knowledge were to be organized on trees. Unlike the combinatory

wheels, these tree-like structures would have enabled the memory in a more direct way. Still, I believe that Yates and Rossi are both correct in including Lull's work in the lineage of the *ars memorativa*. First of all, even though the combinatory wheels may not have enabled the memory, Lull's agenda clearly prioritized memory in a significant manner – one way or another, an aspiring practitioner would have to develop his or her memory substantially to put these techniques into use. If the goal was simply to enumerate all possible combinations of a particular group of concepts, why not simply do so on slips of paper, parchment, or cloth? It is clear that Lull believed that memory was essential to the long-term vision of reducing Christian thought to its logical essence, a process that would be enabled by having practitioners hold the different concepts in their mind in a focused, meditative manner. But more importantly, in understanding the 17<sup>th</sup> century and the emergence of the scientific method, Lull's combinatory wheels would feature significantly in later efforts that aimed to bridge competing schools of thought on memory.

<sup>6</sup>This particular formulation should sound eerily familiar to us – in fact, one way of describing the use of mathematics in the natural sciences would be exactly this, namely, that mathematics is a system for symbolically representing certain relationships inherently present in nature so that we can model nature itself by manipulating these purely symbolic representations. I return to this idea in the final section on the growth of mathematical thought in the 17<sup>th</sup> century.

<sup>7</sup>I am diverging from Yates and Rossi slightly in pairing Leibniz with Ramus. Particularly in comparing Leibniz's efforts to those of Giordano Bruno, I believe it is illustrative to refer to Brunian mnemotechnics as a hybrid mnemonic-Lullian method and Leibniz's method as a hybrid mnemonic-dialectic method. Aside from this comparison, the primary aim of this article does not rest on this particular choice of nomenclature.

<sup>8</sup>An example might help to illustrate Leibniz's views on what a logical structuring of knowledge entailed. One of the primary intermediate goals Leibniz hoped would ultimately lead to a universal calculus was to determine all possible subjects of a given predicate, and conversely, given a subject, to determine all possible predicates. To bridge this encyclopedic, or dialectic, objective with the notion of a symbolic calculus, Leibniz proposed that after fixing a pre-determined ordering of possible predicates, one could associate to each predicate the corresponding prime number (with respect to the fixed ordering). Then, for a subject which satisfied some number of predicates, one could uniquely associate a natural number by taking the product of the corresponding prime numbers. See, for example, [5; pp.178-179].

<sup>9</sup>Stephen Wolfram discusses Leibniz's preoccupation with notation in the beautiful essay, "Dropping in on Gottfried Leibniz," <http://blog.stephenwolfram.com/2013/05/dropping-in-on-gottfried-leibniz>.

<sup>10</sup>Schenkel was a 17<sup>th</sup> century memory theorist.

<sup>11</sup>It should be clear to the reader that I have chosen to focus my attention in this article entirely on developments that took place in Europe. I do not mean to imply that significant efforts in the art of memory did not take place elsewhere. Indeed, there is an incredibly rich tradition of the memory arts in India and it would be an important line of investigation to determine whether in the many scholastic traditions that have emerged across the world, if there have been developments that paralleled those that unfolded in Western Europe, particularly with regard to mnemonic techniques and methods involving symbolic representation, and not simply memorization through repetition.

<sup>12</sup>I wanted to mention here the important work of Mary Carruthers, whose work on the art of memory in Medieval European culture are essential references in the field [19-21]. Many of the earlier innovations of the art of memory I discuss in this article are covered in her work in substantial depth, and I have simply chosen to focus on Yates' and Rossi's treatments because of their closer connection to the events leading up to and during the scientific revolution.

## REFERENCES

- [1] –: Scientific method [Def. 1]. (n.d.).  
Oxford English Dictionary Online, 2013,  
<http://www.oed.com>, retrieved May 16, 2013,
- [2] Nielsen, M.: *Reinventing Discovery: The New Era of Networked Science*.  
Princeton University Press, Princeton 2011,
- [3] –: *Defining the scientific method*. Editorial.  
Nature Methods **6**(4), 237, 2009,  
<http://dx.doi.org/10.1038/nmeth0409-237>,
- [4] Yates, F.A.: *The Art of Memory*.  
Routledge and Kegan Paul, London 1999,
- [5] Rossi, R.: *Logic and the Art of Memory*.  
University of Chicago Press, Chicago 2000,
- [6] Foer, J.: *Moonwalking with Einstein*.  
Penguin, New York, 2011,
- [7] Cicero: *De oratore*.  
II, p.lxxxvi & pp.351-354. Quoted in [4; pp.3-4].
- [8] Magnus, A.: *Opera omnia*.  
Edited by Borgnet, A. Paris, p.97, 1890. Quoted in [4; p.82],
- [9] Quaestio XLIX: *De singulis Prudentiae partibus: articulus I, Utrum memoria sit pars Prudentiae*.  
quoted in [4; p.67],
- [10] Erasmus: *Epistolae*.  
Edited by Allen, P.S. et al. p.IX & pp.29-30. Quoted in [4; p.132],
- [11] Ramus, P.: *De religione Christiana*.  
cd. of Frankfurt, pp.114-115, 1577. Quoted in [4; p.236],
- [12] de Ravenna, P.: *Phoenix sue artificiosa memoria*.  
quoted in [5; p.22].
- [13] –: Algorithm [Def. 2]. (n.d.).  
Oxford English Dictionary Online, 2013,  
<http://www.oed.com>, retrieved May 16, 2013,
- [14] Spampanato, V., ed.: *Documenti della vita di G.B.*  
Florence, pp.84-85, 1933. Quoted in [4; p.200],
- [15] Leibniz, G.W.: *Opusculi et fragmenta inedita de Leibniz*.  
Edited by Couturat, L. Alcan, Paris, pp.27-28, 1903. Quoted in [5; pp.174-175],
- [16] Klein, J.: *Francis Bacon*.  
In Zalta, E.N., ed.: *The Stanford Encyclopedia of Philosophy*. Winter 2012 Edition,
- [17] Bacon, F.: *De augmentis scientiarum*. 5<sup>th</sup> edition.  
Collected and edited by Spedding, J. p.649. Quoted in [4; p.371],
- [18] Descartes, R.: *Cogitationes privatae*.  
Edited by Adam, C. and Tannery, P. *Œuvres de Descartes X*. Vrin, Paris, p.230, 1996. Quoted in [4; pp.373-374],
- [19] Carruthers, M.: *The Book of Memory: A Study of Memory in Medieval Culture*.  
Cambridge University Press, Cambridge, 2008,  
<http://dx.doi.org/10.1017/CBO9781107051126>,
- [20] Carruthers, M. and Ziolkowski, J.M., eds. *The Medieval Craft of Memory: An Anthology of Texts and Pictures*.  
University of Pennsylvania Press, Philadelphia, 2004,
- [21] Carruthers, M.: *The Craft of Thought: Meditation, Rhetoric, and the Making of Images, 400-1200*.  
Cambridge University Press, Cambridge, 1998,

- [22] Wigner, E.P.: *On the unreasonable effectiveness of mathematics in the natural sciences*.  
Communications on Pure and Applied Mathematics **13**(1), 1-14, 1960,  
<http://dx.doi.org/10.1002/cpa.3160130102>,
- [23] Leibniz, G.W.: *Sämtliche Schriften und Briefe herausgegeben von der Preussischen Akademie der Wissenschaften*.  
Edited by Ritter, P. et al. Darmstadt, I & ii & pp.167-169, 1923. Quoted in [5; p.191].

# RECONSIDERING WRITTEN LANGUAGE

Gopal P. Sarma\*

Emory University  
Atlanta, U.S.A.

DOI: 10.7906/indecs.13.3.5  
Regular article

Received: 27 May 2015.  
Accepted: 23 June 2015.

## ABSTRACT

A number of elite thinkers in Europe during the 16<sup>th</sup> and 17<sup>th</sup> centuries pursued an agenda which historian Paolo Rossi calls the “quest for a universal language,” a quest which was deeply interwoven with the emergence of the scientific method. From a modern perspective, one of the many surprising aspects of these efforts is that they relied on a diverse array of memorization techniques as foundational elements. In the case of Leibniz’s universal calculus, the ultimate vision was to create a pictorial language that could be learned by anyone in a matter of weeks and which would contain within it a symbolic representation of all domains of contemporary thought, ranging from the natural sciences, to theology, to law. In this brief article, I explore why this agenda might have been appealing to thinkers of this era by examining ancient and modern memory feats. As a thought experiment, I suggest that a society built entirely upon memorization might be less limited than we might otherwise imagine, and furthermore, that cultural norms discouraging the use of written language might have had implications for the development of scientific methodology. Viewed in this light, the efforts of Leibniz and others seem significantly less surprising. I close with some general observations about cross-cultural origins of scientific thought.

## KEY WORDS

scientific method, scientific revolution, written language, *characteristica universalis*, universal calculus

## CLASSIFICATION

JEL: B19, O31  
PACS: 01.

\*Corresponding author, *η*: gopalsarma@outlook.com; +1 (404) 727-6123;  
201 Dowman Drive, Atlanta, Georgia 30322 USA

## INTRODUCTION

The growth of the scientific method in 16<sup>th</sup> and 17<sup>th</sup> century Europe took place amidst an unusual cultural milieu in which schools of thought on memory and memorization were at the forefront of intellectual developments. During this time period, memory was viewed not only as a set of practices for remembering, but also as a foundational methodology for structuring knowledge [1-3].

Whereas in modern times memorization is often thought of as being antithetical to conceptual thinking, Renaissance and Enlightenment era European culture gave rise to significant innovation in the techniques and applications of memorization. Building on ideas originating in ancient Greece and Rome, European intellectuals from several centuries ago developed highly sophisticated visualization techniques that allowed them to recall massive amounts of information with astonishing accuracy. And in the minds of some of the greatest scientific and philosophical innovators of the time period, notably Francis Bacon, Renes Descartes, and Gottfried Leibniz, as well as other lesser known thinkers, these memorization techniques formed the basic set of concepts around which many thoughts about scientific methodology were formed.

The core premise of this argument – that memory and memorization played a critical role in the emergence of scientific culture – is that the development of the scientific method was not a discrete event. Principled reasoning and systematic investigation have always been part of human society, but during this time period, they became distilled and adopted at an institutional level. How did this transformation take place and why had it not before? In a sense, what forms scientific reasoning consists of a complex and rich set of cultural norms, but which at its core, is rather mundane. By nature, people procrastinate, cut corners, and are not systematic in their efforts. Scientific reasoning might be described as a collection of systematic efforts conducted in a context that prioritizes an attempt to uncover the basic principles governing the behavior of the natural world. How then did the adoption of a systematic approach to knowledge accelerate during the time period traditionally associated with the scientific revolution?

In a previous article, I argued that the art of memory provided an inspiring vision of what could be accomplished with a systematic approach to knowledge [3]. While the basic notions of scientific reasoning may not be abstract, they are not necessarily easy to put into practice, and furthermore, it is not always clear what precise steps would even constitute reasonable next actions for an institution attempting to transform itself into a more rigorous, long-term research establishment. In the absence of a critical mass of scientific accomplishments, major intellectual shifts would have also been difficult to justify. The notion of a journal system and peer review seem elementary to us, but these ideas are far from obvious and required real momentum to become part of the bedrock of institutional practice. At this critical juncture in human history, the art of memory provided a clear, tangible vision and concrete motivation for widespread adoption of a more systematic approach to knowledge. On the relationship of Francis Bacon and Renes Descartes' thinking to the art of memory, for instance, Paolo Rossi writes:

In the works of Descartes and Bacon there is evidence of a direct knowledge of sixteenth-century works on mnemonics. Bacon often mentions 'collection of loci', 'syntaxes' and 'artificial memory', and makes explicit references to the 'doctrine of loci', the 'collocation of images' and the Lullian 'typocosmia'. Descartes, who makes far fewer explicit references and generally refrains from quotation, nonetheless mentions having read Schenkel's *Ars memorativa*, and often refers to the *ars memoriae* and the role played by 'sensible images' in

the representation of intellectual concepts. He also refers to the Lullian idea of the ‘chain of the sciences’ (*catena scientiarum*), mentions his particular interest in one of Lull’s anonymous followers, and asked his friend Isaac Beeckmann about the publication of Agrippa’s Lullist works, and discussed with him the significance and potential of the Lullian art. The themes of the *ars memorativa* and *ars combinatoria* seem to have exercised a significant influence on the thought of Bacon and the young Descartes, and some of these themes can be related directly to Bacon’s conception of a ‘new logic’ and Descartes’s ‘new method’ of philosophizing [2; p.103].

And yet, despite the intense interest of these visionary thinkers, the most ambitious paradigms of this time period did not materialize. It is difficult to imagine that Camillo’s memory theater – an effort to create systematized, physical figurines representing a taxonomy of all knowledge – Bruno’s combinatory wheels, or Leibniz’s universal calculus would have delivered on their promises. But in spite of their failures, these efforts shared the common feature of representing a systematic approach to knowledge. The art of memory and efforts to develop a universal language provided a vision that helped leading intellectuals to clarify and frame their thoughts and to catalyze the adoption of scientific reasoning on a large scale. It created a philosophical context which allowed the more standard and widely discussed factors – patronage, the rise of a journal system, an explicitly articulated notion of hypothesis-driven investigation, etc. – to take root at an institutional level.

## MEMORIZATION AND WRITTEN LANGUAGE

In modern times, there is a standard refrain about the role of memory and memorization in previous eras, namely that in the absence of convenient methods for writing, a trained memory would have been practically significant in ways that it is not today. But something noteworthy about the European schools of thought on memory is that they continued to exist several centuries after the development of the printing press. In examining the culture of this time period, it is clear that memory was viewed in a very different light than by thinkers of our own era. Most broadly, memory was thought of by many 16<sup>th</sup> and 17<sup>th</sup> century intellectuals as being a foundational methodology for structuring knowledge. Furthermore, a few key individuals, notably Gottfried Leibniz, Giordano Bruno, and Renes Descates, conceived of memory as being intimately related to developing symbolic means for representing scientific concepts. For Leibniz, the mnemonic method was a critical element of his vision to develop a universal calculus, a symbolic language to represent and eliminate logical contradictions from the entirety of human knowledge.

While the ambitions of creating a symbolic representation of knowledge is understandable from a modern perspective – certainly, mathematics itself is a symbolic language and has proven itself to be valuable in a wide variety of subjects – what is perplexing about this agenda is the emphasis on memorization. Why was this perspective particularly appealing to thinkers of this time period, when in contemporary society, many practical technological tools – from address books, to electronic calendars, to smart phones – are specifically aimed at relieving our minds from the burden of memory?

With the aim of elucidating the mindset of the time-period, I will suggest the following counterfactual thought experiment. What would a world look like in which efforts analogous to those of Leibniz and his contemporaries thoroughly penetrated all sectors of society? What would be the consequences for a society that relied on memorization for both daily and intellectual tasks, rather than written language?

To imagine this possibility, we can simply return back in time to several thousand years ago before writing systems were widely adopted. In particular, it seems as though, around the

time when written language came into usage, strong cultural norms *discouraging its use* might have precipitated analogous developments to those that took place in 16<sup>th</sup> and 17<sup>th</sup> century Europe. Whereas several millennia had passed between the earliest use of mnemonic techniques in ancient Greece and their flowering in Renaissance and post-Renaissance Europe, a society which resisted the adoption of written language might have seen a comparable developmental trajectory – out of necessity – compressed into a much shorter time period. The result might have been an intellectual explosion, a scientific revolution that is, but of a very different kind than what ultimately came to pass in Europe several millennia later.

Rather than being merely a counter-factual historical curiosity, there is reason to believe that such a set of cultural norms might have actually emerged. The critical observation is that written language – unlike its spoken counterpart – was an *invention*, and like with all inventions, would have been met with some amount of resistance. Some of the animosity towards the written word came from highly informed, intellectual, and influential leaders, and were this resistance to have achieved a critical mass, the adoption of written language might have slowed considerably. The following beautiful passage, from Plato's *Phaedrus*, illustrates this point:

I heard, then, that at Naucratis, in Egypt, was one of the ancient gods of that country, the one whose sacred bird is called the ibis, and the name of the god himself was Theuth. He it was who invented numbers and arithmetic and geometry and astronomy, also draughts and dice, and, most important of all, letters. Now the king of all Egypt at that time was the god Thamus, who lived in a great city of the upper region, which the Greeks call the Egyptian Thebes, and they call the god himself Ammon. To him came Theuth to show his inventions, saying that they ought to be imparted to the other Egyptians. But Thamus asked what use there was in each, and as Theuth enumerated their uses, expressed praise or blame of the various arts which it would take too long to repeat; but when they came to letters, 'This invention, O king,' said Theuth, 'will make the Egyptians wiser and will improve their memories; for it is an elixir of memory and wisdom that I have discovered.' But Thamus replied, 'Most ingenious Theuth, one man has the ability to beget arts, but the ability to judge of their usefulness or harmfulness to their users belongs to another; and now you, who are the father of letters, have been led by your affection to ascribe to them a power the opposite of that which they really possess. For this invention will produce forgetfulness in the minds of those who learn to use it, because they will not practise their memory. Their trust in writing, produced by external characters which are not part of themselves will discourage the use of their own memory within them. You have invented an elixir not of memory but of reminding; and you offer your pupils the appearance of wisdom, not true wisdom, for they will read many things without instruction and will therefore seem to know many things, when they are for the most part ignorant and hard to get along with, since they are not wise, but only appear wise.'

If skepticism of influential leaders such as King Thamus were more widespread and if opposition to written language had reached a critical mass, what might the resulting society have looked like? We would have to imagine that there would be a division of labor entirely devoted to the maintenance of different types of knowledge. Differences in memory that might appear to be extremely subtle to us would be brought to the forefront. There might be groups of people responsible for maintaining long-term knowledge – say, related to agriculture and medicine, or literature even – and others responsible for knowledge that is overturned more quickly, for example, the inventory of vendors in a marketplace.

Furthermore, it seems quite likely that the need to maintain all knowledge in memory would have created a natural and organic selective pressure towards infrastructural simplicity, particularly in the construction of legal and political systems.

To further lend evidence for the plausibility of such a society, I have listed in Table 1 various memory feats that have been performed in ancient and in modern times. The great epics from Greece and India, the Iliad and the Mahabharata, were largely carried down in an oral tradition and were unlikely to have been written down for a long period after they were composed. These were almost certainly memorized using the standard techniques of repetition, although it is worth mentioning that the meter and poetic structure of these epics were significant aids to the memory as well. In our own era, the art of memory has largely been confined to the domain of a rather peculiar set of “memory competitions”<sup>1</sup> in which participants are challenged with a broad array of memorization tasks that seem almost superhuman to those unfamiliar with mnemonic techniques. In the context of the present article, I mention these competitions to give an example of the highly diverse kinds of information that practitioners have developed specialized techniques to remember. For a society built upon memorization, techniques such as these would have no doubt been commonplace – as they were during the Middle Ages, the Renaissance, and the Enlightenment – and one might imagine that for the sake of redundancy and error correction, there would have been social norms encouraging different groups to use different techniques to maintain a given body of information. The ancient art of memory and its recent reincarnation in the form of the competitive memory circuit suggest that both long-term, slowly evolving knowledge, as well as information that is rapidly overturned might have been recorded, processed, re-evaluated, and disseminated without the use of written language.

**Table 1.** Ancient and modern memory feats [5-8].

<b>Task</b>	<b>Size of Text / Time</b>	<b>Techniques Utilized</b>
Iliad	~ 15 000 lines	Repetition / poetic structure
Mahabharata	~ 200 000 lines	Repetition / poetic structure
Speeches of Cicero <sup>2</sup>	Varied	Mnemonic method / repetition
One hour cards <sup>3</sup>	1456 (28 decks)	Mnemonic method
One hour numbers <sup>4</sup>	2660 digits	Mnemonic method
Speed cards (1 deck) <sup>5</sup>	21,19 s	Mnemonic method
Historic dates (5 min) <sup>6</sup>	120	Mnemonic method

It is also worth noting that the consistent usage of mnemonic techniques would have had a measurable impact on brain function and neural development as well. It has been demonstrated, for example, that in contrast to memorization by ordinary repetition, use of the mnemonic method activates those regions of the brain otherwise responsible for spatial navigation [9]. Unsurprisingly, this is due to the fact that the mnemonic method relies explicitly on the use of spatially located images to form memories. Indeed, it is interesting to note that during the Enlightenment, one would have been able to distinguish between disciples of the different schools of thought on memory simply via an fMRI! That is, whereas the dialectic method of Petrus Ramus and his followers would not have given rise to a strong response in the visual regions of the brain, we would expect strong signals from parts of the visual cortex in those practitioners of the mnemonic method, as well as Bruno and Leibniz’s hybrid methods. For a society whose very operational foundation was built upon these techniques, one would expect systematic deviations in cognitive organization from an otherwise normal population. One wonders if these cognitive differences in the capacity to generate intense imagery might have had derivative effects on creativity as well.

While I have presented the above set of circumstances as a counter-factual historical thought experiment, it is worth considering that a scenario such as this one is what Leibniz anticipated would be the societal consequences of the universal calculus, as illustrated in the following excerpt:

My invention contains all the functions of reason: it is a judge for controversies; an interpreter of notions; a scale for weighing probabilities; a compass which guides us through the ocean of experience; an inventory of things; a table of thoughts; a microscope for scrutinizing things close at hand; an innocent magic; a non-chimerical cabala; a writing which everyone can read in his own language; and finally a language which can be learnt in a few weeks, traveling swiftly across the world, carrying the true religion with it, wherever it goes [10].

Aside from the colorful description of the potential impact of this “invention,” what is most striking is the astonishing time frame on which Leibniz believed this language could be acquired. For a person to learn a foreign language in a few weeks would be a superhuman feat, and Leibniz was conceiving of a language that would be acquired on such a time scale which contained not only means for communicating ordinary concepts, but the entirety of human knowledge. In other words, Leibniz was imagining a form of communication that was far superior to ordinary written language. Ultimately, it was hoped that by taking inspiration from the geometric and pictorial aspects of the Chinese alphabet and Egyptian hieroglyphics, and fusing them with the memory-aiding power of the mnemonic method and related memorization techniques, a rigorous process of “notation engineering” would allow for the creation of a symbolic system that transparently represented the factual content as well as the logical structure of the entirety of human thought. It hardly needs to be stated that this vision did not materialize, but it is a striking and original vision and one worth contemplating.

## **DISCUSSION**

If we accept the possibility of a society built upon memorization, and furthermore, that such a set of cultural norms might have had consequences for the development of scientific methodology and scientific culture, we would also have to accept the conclusion that a scientific revolution in such circumstances would have looked very different from what took place in Europe during the 16<sup>th</sup> and 17<sup>th</sup> centuries. For instance, while we traditionally associate the scientific method with hypothesis-driven investigation, one of the primary innovations of the 17<sup>th</sup> century, and which subsequently formed the bedrock of the physical sciences, was Newton and Leibniz’s infinitesimal calculus – on its own, a strictly mathematical theory that is only incidentally related to experimental science and hypothesis-driven investigation. If a scientific revolution were to have taken place at a very different place and time in human history, what might it have actually looked like? What would have been the topics that received the most attention and what would the consequences have been? Forcing ourselves to reason explicitly about such unusual circumstances might help to develop more precise models for what we mean by scientific reasoning, and help to disentangle larger principles from the specific set of historical circumstances in which those principles emerged.

Consider, for example, the recent interest in a scientific study of yoga, Tai Chi, and the meditative practices of the Tibetan Buddhists (see for example [11-13]). Each of these ancient traditions is of sufficient intellectual depth to have drawn the attention of modern scientists. And yet these practices came about from places and during time periods where there were no scientific journals, no means of rapid, widespread communication and double blind peer review, and without the infrastructure of the modern research university. Clearly

then, these early Indian, Chinese, and Tibetan thinkers and experimenters developed a highly sophisticated scientific culture in the absence of many common features of our own scientific world. What would have happened if these efforts were to have been scaled up even more dramatically centuries and millennia ago? What would have the resulting societies and intellectual culture looked like?

In examining the astonishing developments taking place in the contemporary scientific world (see for example [14, 15]), there may be real value in encouraging the writing of counterfactual scientific histories (or at the very least thinking about them), and in particular, asking the question of what a scientific revolution might have looked like at different places and times in human history. The intuition gained from such exercises could prove useful in developing policy recommendations for and guiding the evolution of younger scientific institutions, particularly in the developing world, where there is likely to be more freedom and opportunity to influence the development of institutional culture.

## ACKNOWLEDGEMENTS

I would like to thank Doug Bemis and Aaswath Raman for insightful discussions and critical reading of the manuscript.

## REMARKS

<sup>1</sup>Joshua Foer's memoir *Moonwalking with Einstein* is a beautiful first person account of the history, techniques, and personalities of the "competitive memory circuit" [8]. Perhaps one of the primary lessons of Foer's book relevant to this article is that most of the participants in the various international memory competitions do not claim to have strong natural memories. Rather, their memories were highly trained in the specific context of tasks relevant to competitive memory, for example, memorizing a long list of historical dates, or memorizing the order of a deck of cards. This observation lends further plausibility to the idea of an entire society in which such techniques were commonplace and in which written language was discouraged – these techniques would easily have been learned by many, and need not have been restricted to a handful of elites.

<sup>2</sup>For historical reasons, Cicero's speeches have the distinction of having been memorized through a variety of different techniques [1-3].

<sup>3</sup>In one hour cards, participants memorize the order of as many decks of cards as possible in one hour.

<sup>4</sup>In one hour numbers, participants memorize as many decimal digits as possible in one hour.

<sup>5</sup>In speed cards, participants are given a single shuffled deck of cards (which has been duplicated in another deck) and the task is to memorize the order of all 52 cards in under five minutes.

<sup>6</sup>In historic dates, participants are given a list of fictional events and years and asked to recall as many events as possible after a 5 min training period.

## REFERENCES

- [1] Yates, F.A.: *The Art of Memory*.  
Routledge and Kegan Paul, London 1999,
- [2] Rossi, R.: *Logic and the Art of Memory*.  
University of Chicago Press, Chicago 2000,
- [3] Sarma, G.P.: *The Art of Memory and the Growth of the Scientific Method*.  
Interdisciplinary Description of Complex Systems **13**(3), 373-395, 2015,  
<http://dx.doi.org/10.7906/indecs.13.3.4>,

- [4] Plato: *Phaedrus*.  
Translated by Fowler, H.N. Plato in 12 Volumes. Vol. 9. Harvard University Press, Cambridge & W. Heinemann Ltd., London, Sections 274C-275B, 1925. Quoted in [1; p.38],
- [5] Homer: *Iliad*.  
Edited by Munro, D.B. and Allen, T.W. Oxford University Press, New York, 1920,
- [6] –: *The Mahabharata*.  
Edited by Smith, J.D. Penguin, New York, 2009,
- [7] –: *World Memory Championships*.  
<http://www.world-memory-statistics.com>, retrieved March 21, 2015,
- [8] Foer, J.: *Moonwalking with Einstein*.  
Penguin, New York, 2011,
- [9] Maguire, E.A.; Valentine, E.R.; Wilding, J.M. and Kapur, N.: *Routes to Remembering: The Brains Behind Superior Memory*.  
*Nature Neuroscience* **6**(1), 90-95, 2003,  
<http://dx.doi.org/10.1038/nn988>,
- [10] Leibniz, G.W.: *Sämtliche Schriften und Briefe herausgegeben von der Preussischen Akademie der Wissenschaften*.  
Edited by Ritter, P. et al. Darmstadt, I & ii & pp.167-169, 1923. Quoted in [2; p.191].
- [11] Gard, T., et al.: *Effects of a yoga-based intervention for young adults on quality of life and perceived stress: The potential mediating roles of mindfulness and self-compassion*.  
*The Journal of Positive Psychology* **7**(3), 165-175, 2012,  
<http://dx.doi.org/10.1080/17439760.2012.667144>,
- [12] Kerr, C.E.; Shaw, et al.: *Tactile acuity in experienced Tai Chi practitioners: evidence for use dependent plasticity as an effect of sensory-attentional training*.  
*Experimental Brain Research* **188**(2), 317-322, 2008,  
<http://dx.doi.org/10.1007/s00221-008-1409-6>,
- [13] Harrington, A. and Zajonc, A., eds.: *The Dalai Lama at MIT*.  
Harvard University Press, Cambridge, 2008,
- [14] –: *Defining the scientific method*. Editorial.  
*Nature Methods* **6**(4), 237 2009,  
<http://dx.doi.org/10.1038/nmeth0409-237>,
- [15] Nielsen, M.: *Reinventing Discovery: The New Era of Networked Science*.  
Princeton University Press, Princeton, 2011.

# THE PARADOX OF EXPLOSIVE AND GRADUAL POLICY CHANGE IN POLITICAL REVOLUTIONARY TIMES

Michael Givel\*

Department of Political Science – The University of Oklahoma  
Norman, The United States of America

DOI: 10.7906/indecs.13.3.6  
Regular article

*Received:* 2 July 2015.  
*Accepted:* 21 July 2015.

## ABSTRACT

Many political revolutionary theorists have argued that political revolutionary activity occurs in a dramatic fashion resulting in explosive change in the orientation of established policy regimes resulting in radically new public policy outputs and governmental organizational structures. This research, quantitatively analyzing political revolutions that culminated in the 20<sup>th</sup> century, confirms that short-term political revolutionary activity and the establishment of new policy regimes were few in number. Most successful political revolutionary activities along with new policy regimes were long-term while some political revolutions were not successful. The process of political revolutionary activity to overthrow established policy regimes is a complex phenomenon with political and policy change occurring across widely varying time frames.

## KEY WORDS

political revolution, complexity theory, policy regime

## CLASSIFICATION

JEL: N40

\*Corresponding author, *η*: mgivel@ou.edu; +1 405 325 8878;  
Department of Political Science, The University of Oklahoma, 455 West Lindsey, Room 205  
Norman, Oklahoma 73019, The U.S.A.

## INTRODUCTION

The interplay between revolutionary activity and a political revolution resulting in new policy regimes with radically different policy outputs and governmental organizational structures have been extensively examined in a variety of scholarly studies [1-8]. Policy outputs in this article are government actions or inaction in the form of legislation, executive orders or written mandates, and judicial decisions. A policy regime is defined as governing arrangements among a political coalition or group [9, 10]. In a political revolution, which may be violent or non-violent, a policy regime is overthrown resulting in the enactment of a radically new set of permanent state institutions and policies [11, 12].

One issue of contention in the literature on political revolutions is the length of political revolutions. A number of scholars of political revolutions have postulated that revolutions occur as dramatic and short-term revolutionary change resulting in an explosive change in political power and political structures over a relatively short period of time [13-31]. For instance, Skopcol wrote:

Social revolutions are rapid, basic transformations of a society's state and class structures; and they are accompanied and in part carried through by class-based revolts from below [32].

On the other hand, other scholars have indicated political revolutionary change and policy outputs may occur in a short period, long period, or not at all [13, 14, 33, 34]. In addition, Goldstone also has argued that whether or not and for what duration explosive, dramatic, and short-term revolutionary political change occurs is dependent on:

Where the state remains strong and the opposition is anchored mainly in rural areas, one may see a drawn-out guerilla war. Where the state weakens rapidly, many elites abandon it, and urban groups actively support the opposition, one may see a fairly rapid overthrow of the central authorities [13].

Goldstone also has argued that:

Efforts to change the political regime are based on competing visions of the social order; informal or formal mass mobilization; non-institutional actions including demonstrations, protests, strikes, and violence [19].

In particular, Goldstone's argues that the conditions that initiate a political revolutionary situation include a state in crisis, elites in conflict and in opposition to the state, and a large portion of the population mobilized against the state [19, 35]. In order to ascertain which of these competing theoretical perspectives are accurate, in this paper for all revolutions culminating in the 20<sup>th</sup> century I will examine and analyze all political revolutionary activity resulting in new and permanent policy regimes

## LITERATURE REVIEW

A variety of scholars examining political revolutionary change outside normal political processes have concluded that revolutionary political change can occur in an explosive manner in a short period [13-32, 34, 36]. On the other hand, some scholars of political revolution have also defined revolutionary change as occurring in short periods, long periods, or not at all [13, 14, 33, 34].

Initially in the early 20<sup>th</sup> century, the first generation of revolutionary political theory was conducted by scholars of natural history [18, 19, 37-39]. Here, the focus is on theories based on collective national traits and group or mob psychology based on an unconscious, as

posited by Freudian psychology, “collective mind” [18, 19, 38, 39]. People through a collective mind also initiate group actions such as revolutions. Moreover, people are quite different acting as individuals in contrast to a group or mob. A person’s personality becomes submerged thus allowing the collective mind of the mob to dominate. Based on the unconscious mob mentality, passions, often irrational and exaggerated, rather than the reason of rule is the cause of political revolutions [18, 19, 38, 39].

Beginning in the 1950s, a second generation of scholars of political revolution, following the lead of Talcott Parsons and the structural functionalist school in sociology, utilized policy equilibrium theory to explain societal functions [23-25, 40-42]. A central tenet of the structural functional school is a view of political policymaking systems being in equilibrium and homeostasis or disequilibrium due to shifting balances of power from political demands and conflicts between interest groups [23-25, 40-42]. Political systems under structural functionalism are due to a policymaking system in a holistic model based on important constituent parts. These parts included policy inputs emanating from the total social, cultural, and political environment, conversion of demands into outputs including policy implementation, and feedback [43]. Cultural norms, customs, and political institutions play a primary role in influencing these separate parts in the policy process related to the whole system. Policymaking under structural functionalism often focuses on the process of how a policy is developed. When a system is in extreme disequilibrium, then a punctuated policy reform or even a political revolution can occur [17, 18, 44].

A third generation of scholars of political revolution introduced several new and important variables to understand political revolutions including class and class struggle, the state, international relations and conflicts between states, international capitalist economics, and rural revolts [18, 19, 22, 26, 28, 32, 34]. The focus of this approach was on class conflict and structures. Another key feature of this approach was analyzing class revolution from below. This was manifest in studies of conflicts between peasants or workers and elites who dominated the state. Other foci of this approach was identifying various factors that initiated political revolutions including the role of international capitalist economics and the state as a basis for administrative and coercive power for domestic ruling elites and classes [18, 19, 22, 26, 28, 32, 34].

Goldstone and Foran have argued that since the 1980s a fourth generation of revolutionary political scholarship has emerged [18, 19]. This large and diverse group of scholarship has expanded upon the third generation of political revolution scholarship with a focus on the role of culture, ideology, and leaders in political revolutions [18, 19]. Many fourth generation political revolution scholars have also concluded that there are a wide and complex range of factors that may cause and explain why political revolutions do or do not occur and for what duration [14, 18, 19, 33, 34]. Among these factors that cause political revolutions are the interconnection between state administrative, police and national security structures, ideology, culture, political elites and classes, political leadership, interest groups, mobilizations, and foreign interventions [18, 19]. In addition, Emirbayer and Goodwin argue that important independent factors are the transformation of a political culture including social-psychological and human agency to assess alternative course of action that converts normative views of the world [35].

## **SUMMARY**

With respect to the time period and policy regime change caused by political revolutions, many scholars of political revolution have argued that political revolutions occur in a short period leading to explosive policy change. Some others have asserted that political revolutions may occur in a short period, long period, or not at all. This article will determine

whether political revolutionary activity that attempts to overthrow a political regime are all short-term and explosive in nature leading to dramatic policy output change in the form of a new policy regime.

## **METHODS**

The basic underlying assumptions in measuring the temporal extent of political revolutions that confront political regimes is that political revolutions are based in an “eventful sociology” where the revolutionary process is based on a variety of contingent and complex variables in an ongoing political revolution [45]. What constitutes an eventful sociology can only be determined by historical analyses based on hindsight of revolutionary activity in which a new policy regime was established and not fundamentally changed. The period of 1900 to 1999 was chosen to analyze political revolutionary activity as this represents a suitably long enough recent time period and sample to ascertain trends with respect to whether or not political revolutions occurred explosively and rapidly.

Data on the time period of revolutionary activity and establishment of permanent policy regimes for this analysis was obtained from articles from major newspapers contained in LexisNexis Academic, *New York Times* historical archives, peer reviewed articles obtained from JSTOR, EBSCO, and WorldCat online searches, and from the comprehensive and authoritative encyclopedic history of world revolutions by Goldstone and Defronzo [12, 21, 46, 47].

Measurement of the data from 1990 to 1999 of the eventful sociology of political revolutions will occur in two parts. First, a measurement will be conducted that includes the name of the political revolution, the year the revolution ended, the chronological period in months in which the revolution occurred, the primary groups and opponents involved, and whether the revolution was successful or not. This assessment will determine whether political revolutionary activity is ordinarily short term or not.

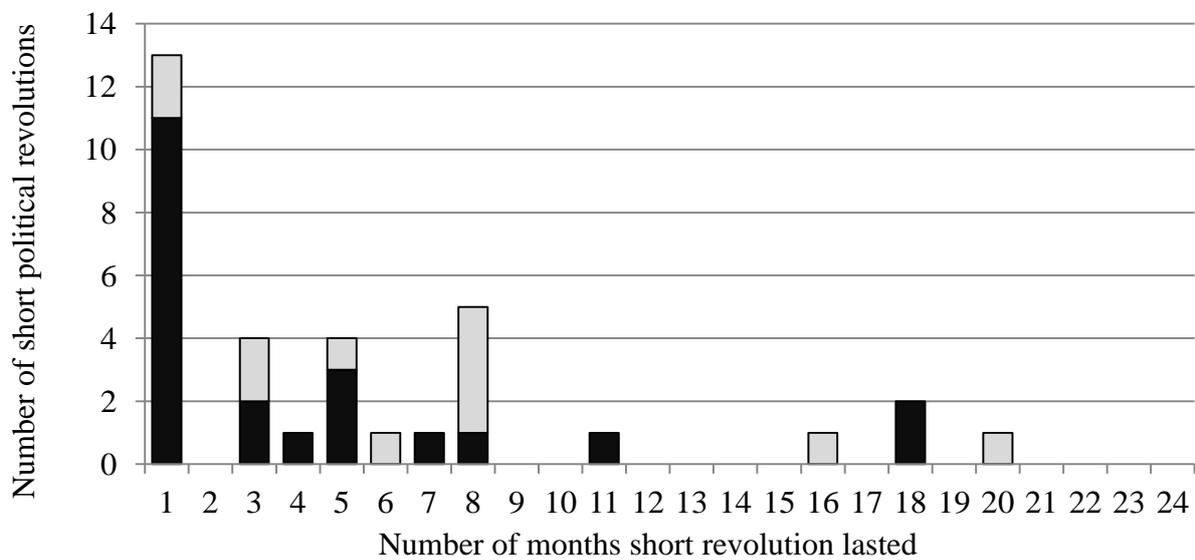
In order to determine whether political revolutionary activity has been rapid and explosive or not, descriptive empirical statistics were generated on the length of each successful or non-successful political revolutionary activity including median, mean, and mode. The time period for each political revolutionary activity was based on the month and year the political revolutionary activity began and ended. For this paper political revolutions are considered relatively short-term if they occurred in 24 months or less. In this paper, two years was utilized as this represents a fairly short period of time while providing a period that does not underestimate short-term political revolutions. Utilizing this approach, I will then determine the nature of the tempo of revolutionary activity that culminated in the 20<sup>th</sup> century.

## **RESULTS**

Appendix provides an overview of each political revolution that ended in the 20<sup>th</sup> century including the number of months of the revolution. As illustrated by Figures 1 and 2, of the 85 instances of political revolutionary activity ending in the 20<sup>th</sup> century, 40 cases or 49 % of political revolutionary activity was short term lasting two years or less while 44 cases or 51 % of political revolutionary activity lasted two or more years and was long term. Additionally, long term revolutionary activities resulting in new policy regimes occurred in 34 out of 45 cases or 75,6 % of the time.

Additionally, with the exception of 21 instances of political revolutionary activity lasting one month or less, all other instances of political revolutionary activity randomly occurred in two instances or less per month. The shortest duration of political revolutionary activity was one month or less and the longest duration manifest in the Indian Independence Movement lasted

751 months. The mode was one month or less as 21 was the greatest number of political revolutionary activities occurring in one time period. The average number of months that political revolutionary activity occurred was 92 months while the median point for political revolutionary activity was 30 months. This indicates that while the number of short-term and long-term political revolutionary activity events were numerically divided almost evenly, the collective length of long-term political revolutionary activity skewed the overall average length and median indicating political revolutionary activity can be quite lengthy in relation to short-term political revolutionary activities. These results confirm that classifications of political revolutions as short-term and dramatic is not accurate in a large majority of cases. In addition, explosive political revolutionary activity and change in a minority of short-term cases to overthrow political regimes was unsuccessful. Finally, the nature of the tempo of political revolutions, overall, represented a complex pattern of time frames in which short and long term political activity occurred.



**Figure 1.** Number of short political revolutions that occurred around world in the 20<sup>th</sup> century, by duration in months. Black bars denote successful, and grey unsuccessful political revolutions.



**Figure 2.** Number of long political revolutions that occurred around world in the 20<sup>th</sup> century, by duration in months. Black bars denote successful, and grey unsuccessful political revolutions.

## DISCUSSION

The basic premise of many political revolutionary theorists is that political revolutionary change occurs in a dramatic and short-term fashion. This research in the area of revolutionary political activities resulting in new policy regimes confirms that in the rare instances when violent and turbulent political revolutionary struggle occurs, most successful political revolutions establishing a new policy regime were not short-term. This description of political revolutions is in line with the arguments of Goldstone and other fourth generation political revolution scholars that political revolutions can be short-term, long-term, or not successful. Political revolutions demonstrate a complex variety of time periods.

In that regard, the study of political revolutionary activity resulting in new policy regimes and public policy outputs and governmental organizational structures needs to come into sync with the conclusion that has now been reached for quite some time by fourth generation scholars of political revolutions. That is, political revolutionary activity linked to the establishment of new policy regimes is a highly varied and even complex matter and often not short-term. Some factors such as culture, ideology, capitalism, corporate actions or the role of classes and elites may have a particularly significant role in causing policy regime changes. Ultimately, business as usual when it comes to comprehending the nature of political revolutionary activity linked to new policy regimes requires analyses that account for complex system behavior as a key feature in ascertaining the nature and scope of political revolutionary actions linked to the policy process.

## APPENDIX

**Table 1.** Political revolutions of the world ending in the 20<sup>th</sup> century. Sources: [12, 21, 46, 47] (continued on pp.411-417).

Name	Years occurred	Time period, month	Revolutionary groups	Opponents	Outcome
Philippine Independence Wars	1872-1910	463	Soldiers, Workers, Catholic Priests	Spanish Colonial Government	did not succeed
Egyptian Revolts	1881-1919	458	Nationalists	Egyptian Government, Ottoman Empire	succeeded
Indian Independence Movement	1886-1947	751	Nationalists, Islamists	British Colonial Government	succeeded
Chinese Boxer Uprising	1900-1901	16	Boxers, Chinese Government	Christians	did not Succeed
Russian Revolution of 1905	1905-1907	31	Workers	Aristocrats, Monarchy, Tsarist Government	did not succeed
Iranian Constitutional Revolution	1906	4	Radical Deputies, Social Democrats, Workers	Iranian Government, Monarchy	succeeded
Turkish Revolution	1908-1922	124	Intellectuals, Workers, Women, Constitutionalists	Turkish Government	succeeded

**Table 1.** Political revolutions of the world ending in the 20<sup>th</sup> century. Sources [12, 21, 46, 47] (continuation from p.410, continued on pp.412-417).

<b>Name</b>	<b>Years occurred</b>	<b>Time period, month</b>	<b>Revolutionary groups</b>	<b>Opponents</b>	<b>Outcome</b>
Mexican Revolution	1910-1940	368	Peasants, Workers, Ranchers, Urban Political Reformers, Socialists, Anarchists	Mexican Government, Armed Forces, Landowners, Employers	Succeeded
Chinese Republican Revolution	1911	3	Faction of Military, Merchants	Chinese Government	Succeeded
Arab "Great Revolt"	1916-1918	29	Hashimites, Nationalists, British	Ottoman Empire	Succeeded
Irish Revolution	1916-1923	82	Irish Nationalists	British Government	Partially succeeded
Russian Revolution of 1917	1917	8	Bolsheviks, Trotskyites	Monarchy, Tsarist Government	Succeeded
German Revolution	1918-1919	8	Communists, Admirals, Sailors	German Government, Social Democrats	Did Not Succeed
Hungarian Revolutions	1918-1919	5	Communists	Hungarian Government	Did Not Succeed
Chinese May Fourth Movement	1919	1	Modernists, Anti-Corruption Advocates, Nationalists, Intellectuals	Chinese Government	Partially Succeeded
Chinese Nationalist Revolution	1919-1927	91	Nationalists	Chinese Government	Succeeded
Italian Fascist Revolution	1919-1945	314	Italian Fascists	Italian Government, Leftists	Succeeded
Chinese Communist Revolution	1921-1949	339	Communists, Intellectuals, Workers, Peasants	Chinese Government	Succeeded
German Nazi Revolution	1923-1945	148	German National Socialists, Fascists	German Government, Leftists	Succeeded
Chinese May Thirtieth Revolution	1925	6	Unions, Workers, Merchants, Students	Chinese Government	Did Not Succeed
Thai Revolution	1932	1	People's Party	Thailand Government	Succeeded
Spanish Civil War	1936-1939	34	Nationalists, Fascists	Spanish Government, Anarchists, Communists, Socialists	Succeeded

**Table 1.** Political revolutions of the world ending in the 20<sup>th</sup> century. Sources [12, 21, 46, 47] (continuation from pp.410-411, continued on pp.413-417).

Name	Years occurred	Time period, month	Revolutionary groups	Opponents	Outcome
Palestinian Anti Colonial Revolt	1936-1939	40	Palestinian Nationalists	British Colonial Government, Zionist Organizations	Did Not Succeed
Pakistani Independence Movement	1940-1947	90	Muslim Nationalists	Indian Government, British Colonial Government	Succeeded
Yugoslav Partisans and Communist Revolution	1941-1948	85	Communists	Yugoslav Government, Ustache (Fascists)	Succeeded
Guatemalan Revolution	1944-1954	122	Students, Peasants, Unions, Dissident Military Officers	Guatemalan Government, Employers, Landowners, Military, Catholic Church, Professionals	Succeeded
Indonesian National Revolution	1945-1950	61	Nationalists, Islamic Revolutionaries, Communists	Netherlands Colonial Government	Succeeded
Venezuelan Democratic Revolution	1945-1958	149	Unions, Dissident Military Officers, Industrialists, Political Parties	Venezuelan Government, Military, Landowners, Employers	Succeeded
Vietnamese Revolution	1945-1975	348	Nationalists, Communists	French and U.S. Colonial Rule	Succeeded
Philippine Huk Revolts	1946-1955	114	Peasant, Tenant Farmers, Laborers, Communists	Philippine Government	Did Not Succeed
Israeli Independence Revolt	1947-1948	26	Zionist Organizations	British Colonial Government, Arab	Succeeded
Madagascar War of Independence	1947-1948	20	Nationalists	French Colonial Government	Did Not Succeed
Ghanaian Independence Movement	1947-1957	117	Nationalists, Unions, Farmers, Populists	Ghana Government, British Colonial Government	Succeeded
Costa Rican Revolution	1948	3	Communist Party	Military, Oligarchy Employers, Costa Rican Government	Succeeded
Malayan Communist Insurgency	1948-1960	136	Communists	British Colonial Government	Did Not Succeed

**Table 1.** Political revolutions of the world ending in the 20<sup>th</sup> century. Sources [12, 21, 46, 47] (continuation from pp.410-412, continued on pp.414-417).

<b>Name</b>	<b>Years occurred</b>	<b>Time period, month</b>	<b>Revolutionary groups</b>	<b>Opponents</b>	<b>Outcome</b>
Columbia's La Violencia	1948-1964	125	Political Parties	Columbian Government, Political Parties, Military	Succeeded
South African Anti-Apartheid Movement	1948-1994	534	African National Congress, Communist Party	South African Government	Succeeded
Korean Civil War	1950-1953	38	North Korean Government, Communists	South Korean Government, Conservatives	Did Not Succeed
Bolivian National Revolution	1952	1	Unions, Workers, Peasants, Socialists	Bolivian Government, Military, Police	Succeeded
Egyptian Revolution	1952	1	Free (Military) Officers, Nationalists	Egyptian Government	Succeeded
Kenyan Mau Mau Movement	1952-1960	89	Sharecroppers, Urban Unemployed, Homeless, Unions	British Colonial Government	Succeeded
Algerian Revolution	1954-1962	93	Nationalists, Religious Activists, Social Democrats, Socialists, and Communists	French Colonial Government	Succeeded
Sudanese Civil War	1955-1972	199	Sudan People's Liberation Movement	Sudan Government, Anglo-Egyptian Condominium	Succeeded
Omani Revolution	1955-1975	267	Tribal Omanis, Marxists, Nationalists	Oman Government	Did Not Succeed
Cuban Revolution	1956-1970	161	Communists, Students	Military, Cuban Government, U.S Companies, Catholic Church	Succeeded
Polish Solidarity Movement	1956-1990	415	Workers, Intellectuals, Writers, Catholic Church	Polish Government, Communists	Succeeded
Iraqi Revolution	1958	1	Military, Free (Military) Officers	Iraq Government, Monarchy	Succeeded
Guinean Independence Movement	1958	1	Nationalists	French Colonial Government	Succeeded

**Table 1.** Political revolutions of the world ending in the 20<sup>th</sup> century. Sources [12, 21, 46, 47] (continuation from pp.410-413, continued on pp.415-417).

<b>Name</b>	<b>Years occurred</b>	<b>Time period, month</b>	<b>Revolutionary groups</b>	<b>Opponents</b>	<b>Outcome</b>
Tibetan Revolt	1959	1	Tibet Nationalists	Chinese Government	Did Not Succeed
Rwandan Civil War	1959-1994	178	Hutus, Peasants	Rwandan Government, Tutsi	Partially Succeeded
Korean Democracy Movement	1960-1998	455	Democracy Movement Workers, Students	South Korean Government	Partially Succeeded
Guinea-Bissau Independence Revolt	1962-1974	152	African Party for the Independence of Guinea and Cape Verde, Communist Party	Portuguese Guinean Colonial Government	Succeeded
Yemeni Revolts	1962-1990	333	Military, Nationalists, Peasants, Marxists	Yemen Government	Succeeded
Eritrean Revolution	1962-1991	346	Eritrean Liberation Front, Communists	Eritrean Government	Succeeded
Syrian Revolution	1963	1	Military, Baathist Socialists, Nasserites, Populists	Syrian Government	Succeeded
Benin Revolutions	1963-1996	391	At First: Military, Marxists; Later: Students, Workers, Unemployed	Benin Government	Succeeded
Zanzibar Revolution	1964	1	Umma Political Party	British Colonial Government	Succeeded
Indonesian Upheaval	1965-1966	7	Military, Communists, Nationalists	Indonesian Government	Succeeded
Chinese Cultural Revolution	1966-1969	36	Red Guards, People's Liberation Army	Bureaucrats, Intellectuals in Chinese Government, Counterrevolutionaries	Succeeded
Zimbabwe Revolt	1966-1980	169	Nationalists	Rhodesian Government	Succeeded
Nigerian Civil War	1967-1970	32	Ethnic Nationalists	Nigerian Government	Did Not Succeed
Cambodian Khmer Rouge	1967-1979	142	Communists, Khmer Rouge	Cambodian Government	Did Not Succeed
Czechoslovakian "Prague Spring"	1968	8	Writers Union, Advocates of "Socialism With a Human Face"	Soviet Union, Warsaw Pact	Did Not Succeed

**Table 1.** Political revolutions of the world ending in the 20<sup>th</sup> century. Sources [12, 21, 46, 47] (continuation from pp.410-414, continued on pp.416-417).

<b>Name</b>	<b>Years occurred</b>	<b>Time period, month</b>	<b>Revolutionary groups</b>	<b>Opponents</b>	<b>Outcome</b>
French Student Revolt	1968	3	Students, Workers	French Government, Communist Party, Unions	Did Not Succeed
Peruvian Revolution	1968-1975	84	Army, Social Reformers, Peasants, Poor	Peruvian Government, Peruvian Navy	Succeeded
Libyan Revolution	1969	1	Military, Nationalists, Socialists	Libyan Government	Succeeded
Chilean Revolution	1970-1973	38	Socialist Party	Military, C.I.A., Landowners, Employers	Succeeded
Bangladeshi War of Independence	1971	11	Independence Movement, Military	Bangladesh Government	Succeeded
Portuguese Revolution	1974-1975	8	Portuguese Military	Portuguese Government	Did Not Succeed
Ethiopian Revolution	1974-1991	141	Military, Ethiopian People's Liberation Front, Ethiopian People's Democratic Revolutionary Front, Anti-Communists	Ethiopian Government	Succeeded
Mozambican Revolution	1974-1994	247	Front for the Liberation of Mozambique, Nationalists	Portuguese Colonial Government	Succeeded
Angolan Revolution	1974-1996	243	National Union for the Total Independence of Angola, Anti-communists	Portuguese Colonial Government, Popular Movement for the Liberation of Angola	Did Not Succeed
East Timorese Independence Movement	1975-1999	181	Students, Catholic Church, Nationalists	Indonesian Government	Succeeded
Afghan Revolution	1978-1995	242	Communists, Islamists, Nationalists	Afghanistan Government	Succeeded
Nicaraguan Revolution	1979	5	Sandinista National Liberation Front	Nicaraguan Government	Succeeded
Iranian Islamic Revolution	1979	1	Nationalists, Business Owners, Professionals, Students, Intellectuals	Iranian Government	Succeeded

**Table 1.** Political revolutions of the world ending in the 20<sup>th</sup> century. Sources [12, 21, 46, 47] (continuation from pp.410-415, continued on p.417).

Name	Years occurred	Time period, month	Revolutionary groups	Opponents	Outcome
Grenada New Jewel Movement	1979-1983	56	Peasants, Youth, Unions, Workers, Women, Large Landowners, Community Organizations, Opposition Politicians	Grenada Government	Succeeded
Philippine People Power Revolution	1986	1	Philippine Citizens	Philippine Government	Succeeded
Haitian Democratic Revolution	1986-1996	119	Professionals, Workers, Unions, Women's Organizations, Youth Groups, Peasants, Neighborhood Organizations, Religious Organizations	Haitian Government, Military	Succeeded
Yugoslav Communist Collapse	1987-1992	53	Nationalists	Yugoslavian Government	Succeeded
Palestinian Intifada	1987-1996	70	Nationalists	Israeli Government	Did Not Succeed
Chinese Tiananmen Square Uprising	1989	3	Pro-Democracy Movement	Chinese Government	Did Not Succeed
Romanian Revolution	1989	1	Democracy Advocates	Romanian Government, Communist Party	Succeeded
Hungarian Anti-Communist Revolutions	1989	5	Party Reformers, Intellectuals, Christian Democrats, Populists, Liberals, Nationalists	Hungarian Government, Communist Party	Succeeded
East German Revolution	1989-1990	5	Perestroika Advocates, Intellectuals, Workers	East German Government, Communist Party	Succeeded
Czechoslovakian "Velvet Revolution" and "Divorce"	1989-1993	40	Writers, Students, Poets, Intellectuals, Nationalists	Czechoslovakian Government, Communist Party	Succeeded
Bulgarian Anti-Communist Revolution	1989-1997	91	Human rights activists, environmentalists, unions, Perestroika advocates, Democracy advocates	Bulgarian Government, Communist Party	Succeeded

**Table 1.** Political revolutions of the world ending in the 20<sup>th</sup> century. Sources [12, 21, 46, 47] (continuation from pp.410-416).

Name	Years occurred	Time period, month	Revolutionary groups	Opponents	Outcome
Baltic Revolutions	1990-1991	18	Democracy Advocates	Latvian, Estonian, and Lithuanian Governments, Communist Party	Succeeded
Albanian Anti-Communist Revolution	1990-1992	18	Intellectuals, Students, Unions	Albanian Government, Communist Party	Succeeded

## REFERENCES

- [1] Hoffertbert, R.I. and Ergruder, U.: *The Penetrability of Policy Systems in a Developing Context*.  
Journal of Public Policy **5**(1), 87-105, 1985,  
<http://dx.doi.org/10.1017/S0143814X00002890>,
- [2] Baumgartner F.R. and Jones, B.D., eds.: *Policy Dynamics*.  
University of Chicago Press, Chicago, 2002,
- [3] Aya, R.: *Theories of Revolution Reconsidered: Contrasting Models of Collective Violence*.  
Theory and Society **8**(1), 39-99, 1979,  
<http://dx.doi.org/10.1007/BF00156400>,
- [4] Janos, A.C.: *Social Science, Communism, and the Dynamics of Political Change*.  
World Politics **44**(1), 81-112, 1991,  
<http://dx.doi.org/10.2307/2010424>,
- [5] Oksenberg, M.C.: *Policy Making Under Mao Tse Tung, 1949-1968*.  
Comparative Politics **3**(3), 323-360, 1971,  
<http://dx.doi.org/10.2307/421471>,
- [6] de Vries, J.: *The Rise and Decline of Nations in Historical Perspective*.  
International Studies Quarterly **27**(1), 11-16, 1983,  
<http://dx.doi.org/10.2307/2600615>,
- [7] Booth, J.A.: *Global Forces and Regime Change: Guatemala in the Central American Context*.  
Journal of Interamerican Studies and World Affairs **42**(4), 59-87, 2000,  
<http://dx.doi.org/10.2307/166342>,
- [8] Steinmetz, G.: *The Colonial State as a Social Field: Ethnographic Capital and Native Policy in the German Overseas Empire before 1914*.  
American Sociological Review **73**(4), 589-612, 2008,  
<http://dx.doi.org/10.1177/000312240807300404>,
- [9] Jochim, A.E. and May, P.J.: *Beyond Subsystems: Policy Regimes and Governance*.  
Policy Studies Journal **38**(2), 303-327, 2010,  
<http://dx.doi.org/10.1111/j.1541-0072.2010.00363.x>,
- [10] Orr, M. and Johnson, V.: *Power and Local Democracy: Clarence Stone and American Political Science*.  
University of Kansas Press, Lawrence, 2008,
- [11] Smith, M.N.: *Rethinking Sovereignty, Rethinking Revolution*.  
Philosophy & Public Affairs **36**(4), 405-440, 2008,  
<http://dx.doi.org/10.1111/j.1088-4963.2008.00147.x>,

- [12] DeFronzo, J.V., ed.: *Revolutionary Movements in World History: From 1750 to the Present, Volume I A-G*. ABC-CLIO, 2006,
- [13] Goldstone, J.: *Revolutions: Theoretical, Comparative, and Historical Studies*. 3<sup>rd</sup> edition. Thomson-Wadsworth, Belmont, 2003,
- [14] Goodwin, J.: *No Other Way Out: States and Revolutionary Movements, 1945-1991*. Cambridge University Press, New York, 2001,  
<http://dx.doi.org/10.1017/CBO9780511812125>,
- [15] Schubert, G.: *Catastrophe Theory, Evolutionary Extinction, and Revolutionary Politics*. In: Somit, A. and Peterson, S., eds.: *The Dynamics of Evolution: The Punctuated Equilibrium Debate in the Natural and Social Sciences*. Cornell University Press, Ithaca, 1989,  
[http://dx.doi.org/10.1016/0140-1750\(87\)90020-0](http://dx.doi.org/10.1016/0140-1750(87)90020-0),
- [16] Artigiani, R.: *Revolution and Evolution: Applying Prigogine's Dissipative Structures Model*.  
*Journal of Social Biological Structures* **10**(3), 249-264, 1987,  
[http://dx.doi.org/10.1016/0140-1750\(87\)90020-0](http://dx.doi.org/10.1016/0140-1750(87)90020-0),
- [17] Goldstone, J.: *Using Quantitative and Qualitative Methods to Forecast Instability*. Special Report No. 204, United States Institute of Peace, 2008,
- [18] Foran, J.: *Theories of Revolution Revisited: Toward a Fourth Generation?*  
*Sociological Theory* **11**(1), 1-20, 1993,  
<http://dx.doi.org/10.2307/201977>,
- [19] Goldstone, J.: *Toward a Fourth Generation of Revolutionary Theory*.  
*American Review of Political Science* **4**, 139-187, 2001,  
<http://dx.doi.org/10.1146/annurev.polisci.4.1.139>,
- [20] Kroeber, C.B.: *Theory and History of Revolution*.  
*Journal of World History* **7**(1), 21-40, 1996,  
<http://dx.doi.org/10.1353/jwh.2005.0056>,
- [21] Goldstone, J.: *The Encyclopedia of Political Revolutions*.  
Congressional Quarterly, Washington, D.C., 1998,
- [22] Moore, B.: *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*.  
Beacon Press, Boston, 1966,
- [23] Johnson, C.: *Revolutionary Change*.  
Stanford University Press, Palo Alto, 1982,
- [24] Gurr, T.: *Persistent Patterns of Repression and Rebellion: Foundations for General Theory of Political Coercion*.  
In: Karns, M., ed.: *Why Rebellions Persist*. Praeger, New York, 1986,
- [25] Huntington, S.: *Political Order in Changing Societies*.  
Yale University Press, New Haven, 1968,
- [26] Paige, J.M.: *Conjuncture, Comparison, and Conditional Theory in Macrosocial Inquiry*.  
*American Journal of Sociology* **105**(3), 781-800, 1999,  
<http://dx.doi.org/10.1086/210360>,
- [27] Trimberger, E.K.: *Revolutions From Above; Military Bureaucrats and Development in Japan, Turkey, Egypt, and Peru*.  
Transaction Books, New Brunswick, 1978,
- [28] Tilly, C.: *From Mobilization to Revolution*.  
Addison-Wesley, Reading, 1978,
- [29] Foran, J.: *The Comparative-Historical Sociology of Third World Social Revolutions: Why A Few Succeed, Why Most Fail*.  
In: Foran, J., ed.: *Theorizing Revolutions*. Routledge, London, 1997,  
<http://dx.doi.org/10.4324/9780203206638>,
- [30] Goldstone, J.: *Revolution and Rebellion in the Early Modern World*.  
University of California Press, Berkeley, 1991,

- [31] Halliday, F.: *Revolution in the Third World: 1945 and After*.  
In: Rice, E.E., ed.: *Revolution and Counterrevolution*. Verso, London, 1991,
- [32] Skopcol T. *States & Social Revolutions*. New York: Cambridge University Press; 2003.
- [33] Halliday, F.: *Counter-Revolution*.  
In: Goldstone, J., ed.: *Revolutions: Theoretical, Comparative, and Historical Studies*.  
Thomson-Wadsworth, Belmont, 2003,
- [34] Wolf, E.: *Peasant Wars of the Twentieth Century*.  
Harper, New York, 1969,
- [35] Emirbayer, M. and Goodwin, J.: *Symbols, Positions, Objects: Toward a New Theory of  
Revolutions and Collective Action*.  
*History and Theory* **35**(3), 358-374, 1996,  
<http://dx.doi.org/10.2307/2505454>,
- [36] Marx, K. and Engels, F.: *Manifesto of the Communist Party*.  
1848,
- [37] Edwards, L.: *The Natural History of Revolution*.  
University of Chicago Press, Chicago, 1927,
- [38] Pettee, G.: *The Process of Revolution*.  
Harper and Row, New York, 1937,
- [39] Bon, G.L.: *The Psychology of Revolution*.  
Batoche Books, 1913,
- [40] Parsons, T.: *Politics and Social Structure*.  
Free Press, New York, 1969,
- [41] Davies, J.: *Toward a Theory of Revolution*.  
*American Sociological Review* **27**, 5-19, 1962,  
<http://dx.doi.org/10.2307/2089714>,
- [42] Smelser, N.J.: *Theory of Collective Behavior*.  
Free Press, New York, 1963,  
<http://dx.doi.org/10.1037/14412-000>,
- [43] Easton, D.: *A Systems Analysis of Political Life*.  
Wiley, New York, 1965,
- [44] Baumgartner, F.R. and Jones, B.D.: *Agendas and Instability in American Politics*. 2<sup>nd</sup> edition.  
University of Chicago Press, Chicago, 2009,  
<http://dx.doi.org/10.7208/chicago/9780226039534.001.0001>,
- [45] Sewell, W.: *Three Temporalities: Toward an Eventful Sociology*.  
In: McDonald, T., ed.: *The Historic Turn in the Human Sciences*. University of Michigan Press,  
Ann Arbor, pp.245-280, 1996,
- [46] DeFronzo, J.V., ed.: *Revolutionary Movements in World History: From 1750 to the  
Present, Volume 2 H-P*.  
ABC-CLIO, 2006,
- [47] DeFronzo, J.V., ed.: *Revolutionary Movements in World History: From 1750 to the  
Present, Volume 3 R-Z*.  
ABC-CLIO, 2006.

# RECENT IMPACTS OF SELECTED DEVELOPMENT INDICATORS ON UNEMPLOYMENT RATE: FOCUSING THE SEE COUNTRIES

Ksenija Dumcic<sup>1,\*</sup>, Vesna Bucevska<sup>2</sup> and Emina Resic<sup>3</sup>

<sup>1</sup>Faculty of Economics and Business University of Zagreb  
Zagreb, Croatia

<sup>2</sup>Faculty of Economics of Ss. Cyril and Methodius University in Skopje  
Skopje, R. Macedonia

<sup>3</sup>The School of Economics and Business University of Sarajevo  
Sarajevo, Bosnia nad Herzegovina

DOI: 10.7906/indecs.13.3.7  
Regular article

Received: 9 June 2015.  
Accepted: 17 July 2015.

## ABSTRACT

The article investigates the relationship between unemployment rate and development indicators: (1) the GDP per capita in Purchasing Power Parities (PPP in current international \$); and (2) the Internet penetration rate, defined as the percentage of Internet users per 100 people. For 34 countries in 2013, only two simple linear regression models based on natural logarithms of data and the Ordinary Least Squares (OLS) estimator appeared to be useful. The simple linear regression *Model 1* shows a negative correlation between the main variable under study  $\ln Y_{UemRate}$  and the regressor  $\ln X_{GDPpc}$ , explaining nearly half of the total variation. The simple linear regression *Model 2* shows a negative correlation between  $\ln Y_{UemRate}$  and  $\ln X_{IntUse}$ , explaining 27 % of the total sum of squares. Regarding clustering of 34 countries based on three variables, the Ward linkage and squared Euclidean distances gave an interesting four-cluster solution. The South-East European (SEE), and especially to the Western Balkan's countries (WBC) are focused. These countries, spread in three clusters, are not homogeneous. Bosnia and Herzegovina and R. Macedonia are with Spain and Greece, all having difficult economic situation. Albania, Montenegro and Serbia are with Bulgaria, Romania and Turkey, all being the SEEC. Croatia is with more developed Italy, Cyprus and Poland, and with less developed Portugal. Central European Slovenia, joined more developed countries of that area, but the most developed European countries comprised a cluster of their own.

## KEY WORDS

unemployment rate, GDP per capita in PPP, Internet penetration rate, multivariate analysis, Western Balkan countries

## CLASSIFICATION

JEL: C51, O12, O57

\*Corresponding author, *η*: kdumcic@efzg.hr; +385 1 2383 363;  
Faculty of Economics and Business, Trg J.F. Kennedyja 6, HR – 10 000 Zagreb, Croatia

## **INTRODUCTION**

In this article the impact of selected economic and Information and Communication Technology (ICT) indicators on unemployment rate in selected European countries is analysed. The aim of the research is to study the relationship of the unemployment rate (the percentage of total labour force), as the dependent variable, and two selected development indicators: the GDP per capita and the Internet penetration rate.

Regarding the geographic scope of the research, the official data availability for variables under study determined the countries for this analysis. The World Bank Data are used [1]. In the focus are the recent changes in the SEE and the WB countries. There are different trends in all three variables from 1990 to 2013 in each of 37 analysed European countries.

Based on statistical data exploration, upon recognising and deleting the outliers, the regression analysis performed for cross-sectional data for 34 countries for 2013.

The goal was to study whether and to what extent do the GDP per capita in PPP in current international \$ ( $X_{GDPpc}$ ), and the Internet penetration rate, as the percentage of Internet users per 100 people ( $X_{IntUse}$ ), impact the total unemployment rate  $Y_{UemRate}$ . For this purpose several regression models, firstly, using the original data, and secondly using the logarithms of data, are developed. The cluster analysis are conducted with the purpose of testing the research hypothesis that the similarities among the SEE (and WB) countries, and among highly developed countries exist.

## **LITERATURE OVERVIEW**

According to European Commission [2] the Labour market conditions in Europe started to improve recently during 2013. With output growth accelerating only slowly, and given the usual lagged response of employment, small net job creation is expected in the short term. Public employment growth is set to remain Intra-euro-area rebalancing is progressing, while private employment growth is still dampened by the remaining scope for firms to adjust working hours. Employment growth in 2014 is expected to be limited, at 0,6 % in the EU and 0,4 % in the euro area. The unemployment rate is thus expected to decrease slightly in 2014 from its very well-known peak in 2013. In 2015, employment growth is set to accelerate to 0,7 % in both areas, resulting in a further slight reduction of unemployment to around 10 % in the EU and 11 % in the euro area. Such a slow decline reflects the gradual recovery but could also indicate a higher prevalence of structural unemployment than in the pre-crisis period. Large differences in labour market performance would persist although unemployment is set to decrease in a large majority of the EU Member States.

According to [3] an analysis of youth unemployment in the Euro Area was investigated. The paper starts by presenting some facts on youth unemployment over the last two decades at the euro area and at each of the European countries level. Over the last few years, youth unemployment has remained at a high level relative to other age groups in most Euro Area countries. The paper finds that there is a positive relationship between the share of young people in the total population and the youth unemployment rate. It is shown that the smaller the percentage of young people in the population, the lower the risk of them being unemployed. At the same time, economic conditions are negatively correlated with the youth unemployment rate. So, the youth unemployment rate increases when the economic situation worsens. Moreover, robust results across the regression scenarios show that higher employment protection and minimum wages imply a higher youth unemployment rate, while active labour market policies (ALMPs) tend to reduce it. The research results also indicate that the increasing share of services employment in total employment is helping to reduce

unemployment among young people. The increase in the youth inactivity rate, which is mainly due to the fact that there are more young people in education, is also connected to the overall decline in youth unemployment. Regarding education, the results indicate that number of years, i.e. the length of education is associated with lower youth unemployment rates. The share of the young population not in school, however, is positively correlated with the unemployment rate. As youth unemployment is subject to certain country-specific features, each country should identify the relevant underlying sources of youth unemployment and react accordingly. European countries' governments can make a positive contribution to the smooth transition of young people from education to the labour market by supporting a well-functioning education system and labour market institutions that do not introduce distortions into the labour market.

The relationship between unemployment and economic growth in Jordan through the implementation of Okun's law is presented in [4]. Time series of annual data from the period 1970-2008 are studied. The relation between unemployment and economic growth obtaining estimates for Okun's coefficient are tested. The study used Augmented Dickey-Fuller (ADF) for unit root, co-integration test and a simple regression between unemployment rate and economic growth. The empirical results reveal that Okun's law cannot be confirmed for Jordan. Thus, it can be suggested that the lack of economic growth does not explain the unemployment problem in Jordan. The author recommended that economic policies related to demand management would not have an important effect in reducing unemployment rate. Accordingly, implementation of economic policies oriented to structural change and reform in the labour market would be more appropriate by policy makers in Jordan.

Although there is a significant literature on the relationship between economic growth and unemployment, effect of economic growth over unemployment varies among the periods and countries. The study given by [5] investigates the economic growth, productivity and unemployment data for seven industrialized countries (G7) between the years of 2000 to 2011. In addition to the mentioned period two sub-periods of 2000-2007 and 2008-2011 in which the effect of global financial crisis was felt most have been analysed. Pre and post crisis periods are compared to each other. The results of this study reveal that while the productivity and economic growth variables have significant and strong effects on the decrease of unemployment in the pre-crisis period, this effect of productivity becomes insignificant and small after the crisis whereas the effect of economic growth as a decreasing effect over unemployment continues and its impact level rises.

Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts through an empirical comparison for the G7 countries is investigated in [6].

There are several recent articles developing and evaluating forecasts of unemployment in European countries. The paper [7] predicted macroeconomic indicators in the Czech Republic using econometric models and exponential smoothing techniques, while [8] evaluated the accuracy and bias of the unemployment rate forecasts suggesting methods of improving the forecasts accuracy. In [9] the forecasts for inflation and unemployment rate based on models using resampling techniques are given.

The performance of unemployment rate predictions in Romania developing strategies to improve the forecasts accuracy are presented in [10]. Voineagu, In [11] the authors forecasted monthly unemployment using econometric smoothing techniques.

According to [12] regarding male and female unemployment trends, historically, women have been more affected by unemployment than men. Recent analysis conducted by Eurostat revealed some interesting trends in unemployment rates by gender in European and non-European countries appeared.

Potential forecasting models suitable for predicting the future values of unemployment rates for male and female based on annual data from 1991 to 2013 in 12 European countries are explored in [13]. A highly developed EU country like Austria, but also others less developed countries, such as, Spain, Greece, both with the highest unemployment rates, Croatia, Portugal, Slovenia, and EU candidates, Bosnia and Herzegovina, Serbia, Turkey, Albania, R. Macedonia and Montenegro are investigated. The results of the empirical analysis showed that the optimal model for forecasting unemployment rate is different for different countries. The paper gives the insights in the most appropriate forecasting methods among regression models and smoothing methods for predicting unemployment rate by gender. Even with the best fitted models the real unemployment rates might be either under- or over-estimated.

Forecasting models for unemployment rate in selected European countries using smoothing methods are evaluated in [14].

## **DATA AND RESEARCH METHODOLOGY**

### **DATA SOURCE AND DEFINITIONS**

The objective of the research is to study the relationship between the unemployment rate, defined according to the International Labour Organisation as percentage of total labour force ( $Y_{UmRate}$ ), determined in the regression analysis as the dependent variable, and two selected development indicators. The first indicator used as the regressor is GDP per capita in PPP in current international \$,  $X_{GDPpc}$ . The second regressor is the indicators called Internet penetration rate, given as the percentage of Internet users (per 100 people),  $X_{IntUse}$ .

According to [15] Eurostat defines an unemployed person as person aged 15-74 without job during the reference week who is available to start work within the next two weeks and who has actively sought employment at some time during the last four weeks. Unemployment rates represent unemployed persons as a percentage of the labour force. The labour force is the total number of people employed and unemployed. Unemployed persons are persons aged 15 to 74 who were: without work during the reference week; currently available for work (i.e. were available for paid employment or self-employment before the end of the two weeks following the reference week); and actively seeking work (i.e. had taken specific steps in the four weeks period ending with the reference week to seek paid employment or self-employment or who found a job to start later i.e. within a period of, at most, three months).

The World Bank Data time series data combined with the Eurostat data for the period from 1990 to 2013 were analysed [12, 15].

### **RECENT DYNAMICS OF UNEMPLOYMENT RATES**

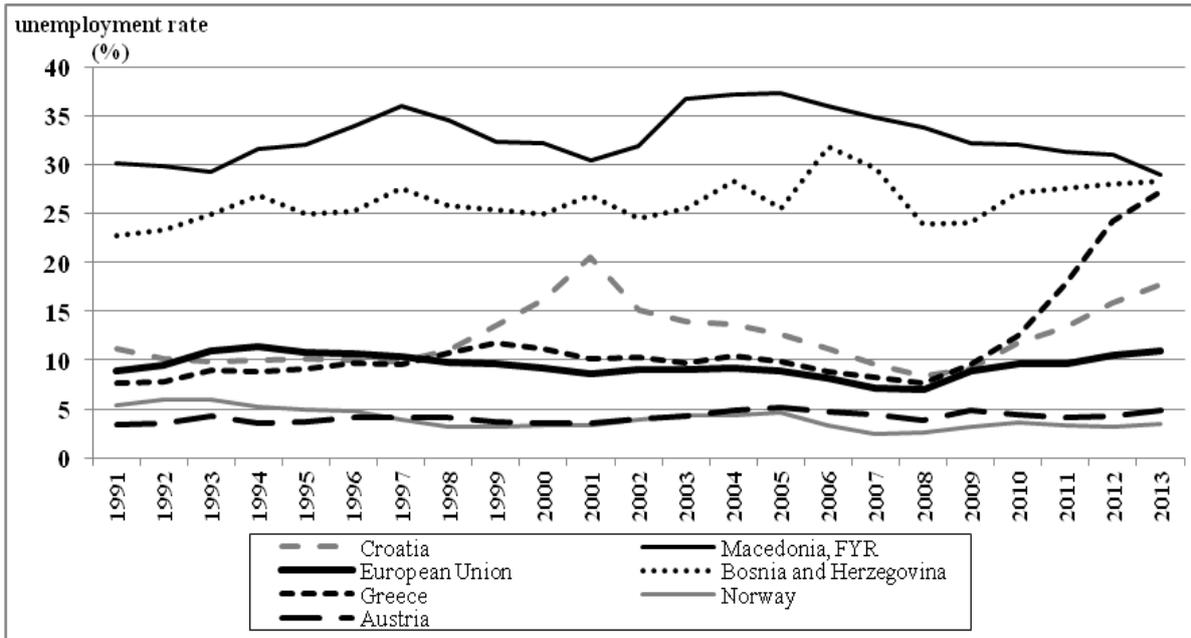
The unemployment rate showed different trends over 37 analysed European countries. Figure 1 presents the lowest and the highest unemployment rates.

After the ILO survey, the total worldwide unemployment rate in 2015 is forecasted to remain unchanged at the level of 5,9 % compared to the previous year, being the highest (12,5 %) in the North Africa, and the lowest (3,9 %) in South Asia.

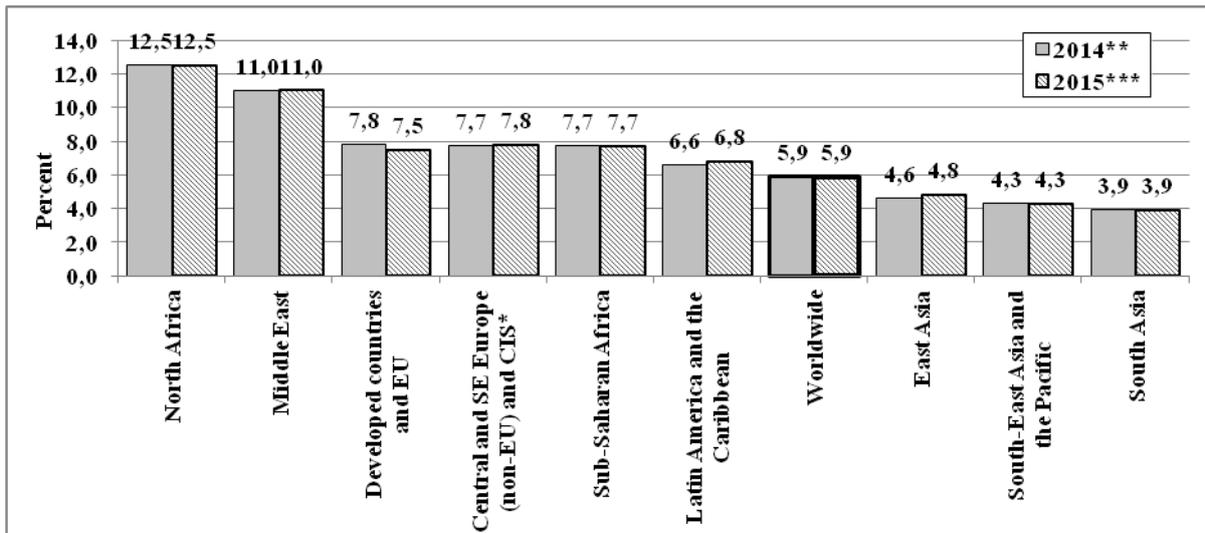
### **UNEMPLOYMENT RATES BY GEOGRAPHY**

Figure 2 shows Unemployment rate in selected world regions 2014 after worldwide surveys conducted by International Labour Organisation in 2014 and 2015.

According to the Eurostat survey data from November 2014 [12], Figure 3 resulted.



**Figure 1.** Unemployment rates in the period 1991 to 2013: the lowest for Norway and Austria, the highest for R. Macedonia, Bosnia and Herzegovina and Greece; and for Croatia and the EU-28 average. Source: Authors’ creation and [16].

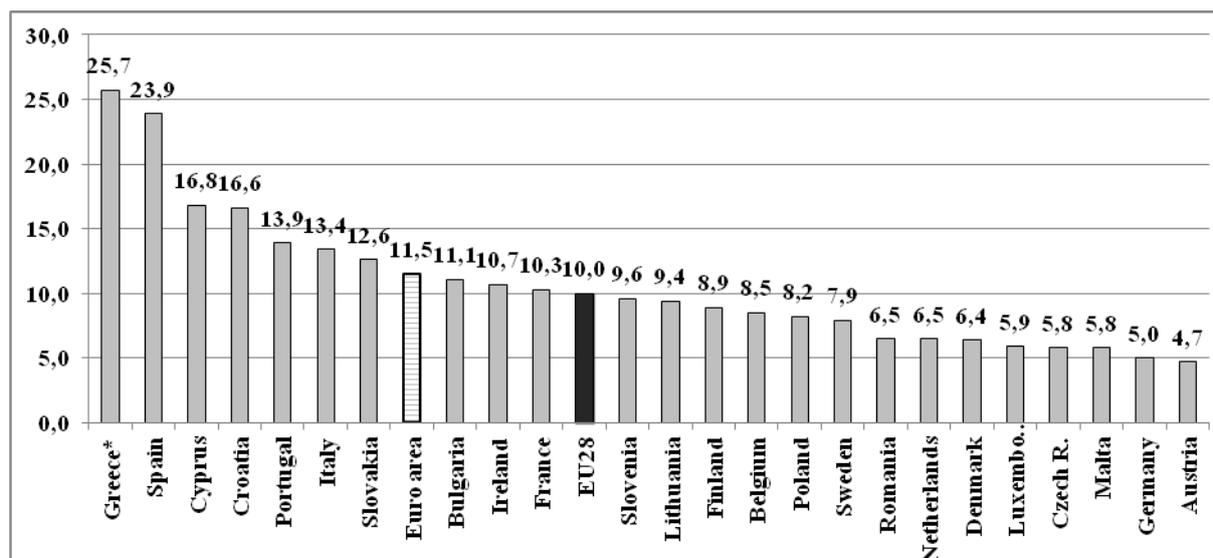


**Figure 2.** Unemployment rate in selected world regions in 2014 and 2015. Source: Authors’ creation and [15]. Notes: \* – the abbreviation CIS stands for Commonwealth of Independent States; \*\* – provisional estimate, \*\*\* – forecast.

### DESCRIPTIVE DATA EXPLORATION

Table 1 shows the descriptive statistical analysis results. In 2013 the outlier for the Luxembourg GDP per capita is indicated with standardized value  $z = 3,52$ . Its standardized value for this variable was even larger in some past years.

After exploration of data variability, data for three highly developed countries Luxembourg, Iceland and Norway are excluded, being the outliers in most of the recent years regarding the GDP per capita. Figure 4 shows multiple Box Plot data for all three variables for 2013, where data for Luxembourg could be noticed as a seriously high outlier. Figure 4 indicates that the distribution of data for  $Y_{UmRate}$  is positively skewed with quite high skewness,  $\alpha_3 = 1,16$ .



**Figure 3.** Unemployment rate in member states of the European Union in November 2014 (seasonally adjusted). Source: Authors' creation and [12]. Note: For Germany, the Netherlands, Austria, Finland and Iceland the trend component instead of the more volatile seasonally adjusted data is used.

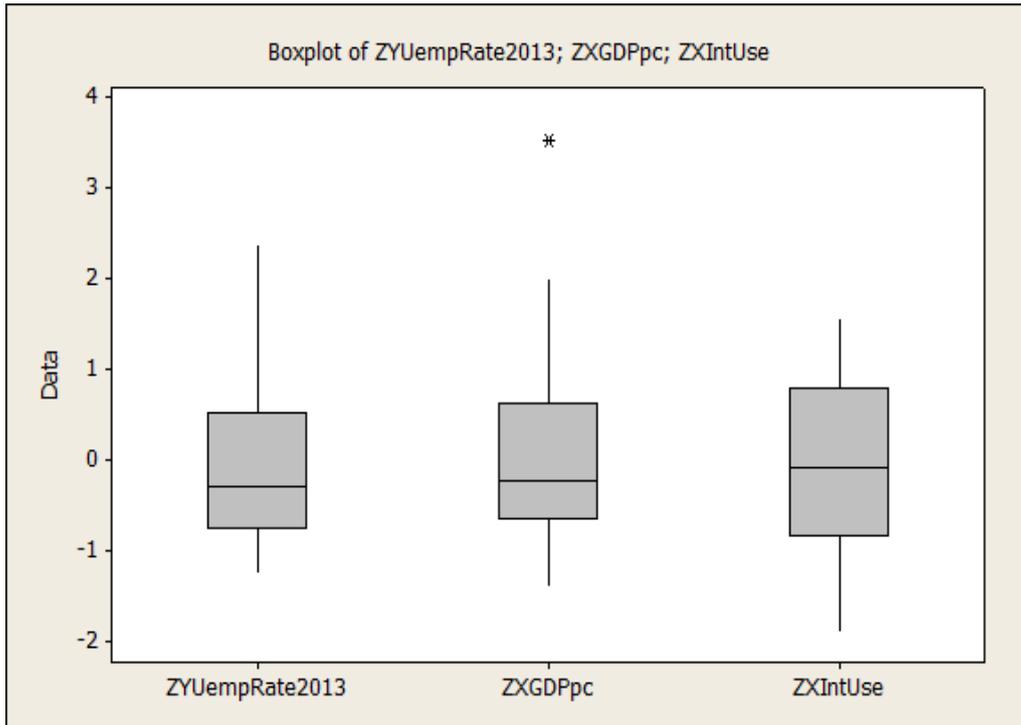
Just for exploration, the scatter diagrams for  $Y_{UemRate}$  and  $X_{GDPpc}$ , and  $Y_{UemRate}$  and  $X_{IntUse}$ , are given in Figure 5. Only weak negative correlation might be seen for both pairs of variables.

Finally, 34 countries remained for further regression analysis: 27 from EU28 countries' data (for all countries but not for the Luxembourg data); 3 official EU candidates (R. Macedonia, Serbia and Turkey); 3 potential EU candidates (Albania, Bosnia and Herzegovina, Montenegro), plus Switzerland. From 12 South-East European (SEE) countries, 11 of them are analysed: Albania, Bulgaria, Bosnia and Herzegovina, Cyprus, Greece, Croatia, R. Macedonia, Montenegro, Romania, Serbia and Turkey. Only Kosovo could not be included because of the lack of data.

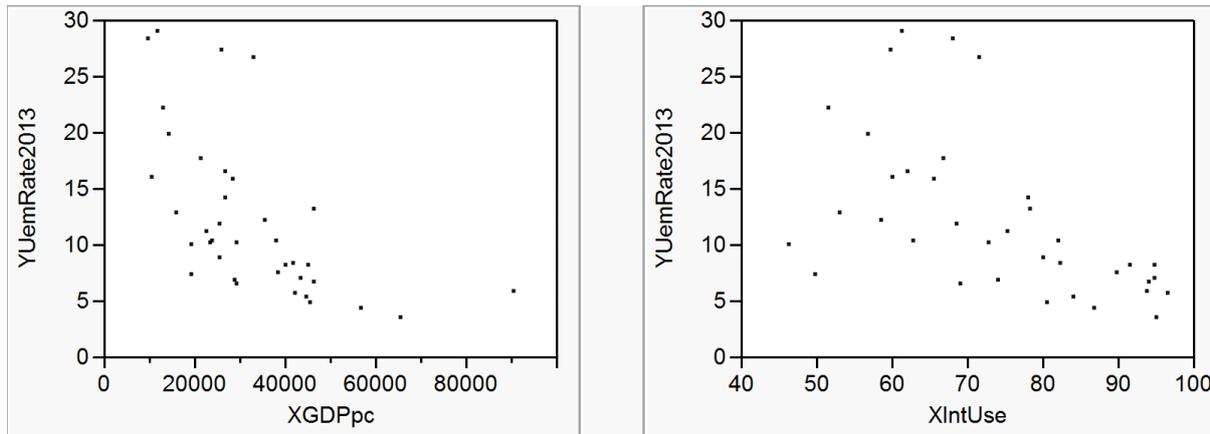
**Table 1.** Descriptive statistics for 2013 data.

Descriptive Measure	$Y_{UemRate}$	$X_{GDPpc}$	$X_{IntUse}$
Mean	12,18	32441,14	73,71
Standard Error	1,16	2706,07	2,38
Median	10,20	28769,86	72,68
Mode	10,40	not defined	not defined
Standard Deviation	7,04	16460,41	14,48
Kurtosis	0,52	3,06	-1,02
Skewness	1,16	1,32	-0,02
Range	25,50	80874,60	50,30
Minimum	3,50	9535,54	46,25
Maximum	29,00	90410,14	96,55
Sum	450,80	1200322,08	2727,26
Count	<b>37</b>	<b>37</b>	<b>37</b>
Coefficient of variation	0,58	0,51	0,20
$Z_{min}$	-1,23	-1,39	-1,90
$Z_{max}$	2,39	3,52*	1,58

\*indicates the outlier for Luxembourg



**Figure 4.** Multiple Box Plot for standardised data for all three variables and 37 countries for 2013. Source: authors’ creation and [16].



**Figure 5.** Scatter plots for pair of variables:  $Y_{UemRate}$  and  $X_{GDPpc}$ ; and  $Y_{UemRate}$  and  $X_{IntUse}$ ,  $n = 37$  countries, data for 2013. Source: authors’ creation and [16].

Considering the South-East European (SEE) region, from altogether 12 SEE countries, even 11 of them are included into the research. There are data for five EU member states that fall into the SEE: Bulgaria, Croatia, Cyprus, Greece and Romania, as well as data for the rest of the SEE countries: three official EU candidates (R. Macedonia, Serbia and Turkey all belonging to the Western Balkans), as well as three potential EU candidates (Albania, Bosnia and Herzegovina and Montenegro), are analysed. Slovenia, as the part of the West-Central Europe, is not included into the SEE region.

Altogether six countries, Albania, Bosnia and Herzegovina, R. Macedonia, Kosovo, Montenegro and Serbia belong to the WB region. Sometimes, according to [17], Croatia is added to the WB group of countries in the research reports. Kosovo’s data are not available, so, it could not be included into the analysis shown in this article.

## RESEARCH RESULTS AND FINDINGS

### REGRESSION MODELLING

The aim of the research is to study the relationship of the unemployment rate (percentage of total labour force, modelled as the International Labour Organisation estimate, compare to [15]), as the dependent variable, and two selected development indicators. The regression analysis using cross-sectional data for 2013 was conducted with the purpose to investigate to what extent the GDP per capita in PPP (in current international \$) and the Internet penetration rate, i.e. percentage of Internet users per 100 people, impact the total unemployment rate. The firstly developed multiple linear regression model was not statistically significant in the variable  $X_{InteUse}$ , so, it was not found to be appropriate for use. In addition, two simple linear regression models were evaluated, both having the heteroskedasticity problems. Therefore, the logarithmic transformation for the variables is suggested. All possible regressions with logarithms of data were investigated, but only two simple logarithmic regression models are shown to be useful, with no violation of model assumptions.

The regression models are built for 34 European countries, with the population model given as follows:

$$\ln y = \beta_0 + \beta_1 \ln x_1 + \varepsilon. \quad (1)$$

For estimating the regression parameters from the model (1) the OLS estimator was applied.

The regression model with estimated parameters is:

$$\ln \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \ln x_1. \quad (2)$$

The estimated Model 1 for regressing the  $\ln Y_{UemRate}$  on  $\ln X_{GDPpc}$  for 34 in 2013 is:

$$\begin{aligned} \ln \hat{Y}_{UemRate2013} = 10,0141 - \mathbf{0,7448} \cdot \ln X_{GDPpc} \quad n = 34 \quad R^2 = 0,4587 \quad \bar{R}^2 = 0,4418 \\ (1,4597) \quad (0,1430) \quad \hat{\sigma} = 0,3832 \quad \hat{V} = 15,83\% \quad DW = 2,386 \quad (3) \\ R = 0,6773 \quad F = 27,122 \end{aligned}$$

All the assumptions of the regression Model 1 are filled. According the Breusch-Godfrey Serial Correlation LM test (p-value equal to 0,4150), there is no autocorrelation problem on the significance level of 5 %. Moreover, the Jarque-Bera test (p-value equal to 0,2386) shows that error terms are normally distributed on significance level of 5 %. In order to check the analysed regression model for heteroskedasticity, the Breusch-Pagan-Godfrey test shows at the 5 % significance level, with p-value equal to 0,7312, the variance is stable. The individual t-test of significance shows that the independent variable (p-value < 0,0000) is statistically significant at 5 % significance level.

After the Model 1 is estimated, it might be concluded that if the *GDP per capita*,  $X_{GDPpc}$ , would increase by 1 %, the regression value of  $Y_{UemRate2013}$  would decrease by 0,7448 %. The estimated model explains 45,87 % of the total sum of squares, and the regression coefficient of variation is 15,83 %, so the estimated model might be considered as a representative one.

Further, the estimated Model 2 for regressing the  $\ln Y_{UemRate2013}$  on  $\ln X_{IntUse}$  for 34 countries in 2013 is:

$$\begin{aligned} \ln \hat{Y}_{UemRate2013} = 8,2864 - \mathbf{1,3783} \ln X_{IntUse} \quad n = 34 \quad R^2 = 0,2712 \quad \bar{R}^2 = 0,2484 \\ (1,7017) \quad (0,3994) \quad \hat{\sigma} = 0,4447 \quad \hat{V} = 18,37\% \quad DW = 2,495 \quad (4) \\ R = 0,5208 \quad F = 11,908 \end{aligned}$$

If the Internet Penetration Rate,  $X_{IntUse}$ , would increase by 1 %, the regression value of  $Y_{UemRate2013}$  would decrease by 1,3783 %. The estimated model explains 27,12 % of the total variation, and the regression coefficient of variation is 18,37 %. Therefore, the estimated model is a representative. The regression model assumptions are not violated. According the Breusch-Godfrey Serial Correlation LM test (p-value equal to 0,2062), there is no autocorrelation problem on the significance level of 5 %. The Jarque-Bera test (p-value equal to 0,9566) shows that error terms are normally distributed at 5 % significance level. In order to check the analysed regression model for heteroskedasticity, the White test was conducted (p-value equal to 0,1598) and shown that at the 5% significance level, the variance is stable. Individual t-test of significance shows that the independent variable (p-value < 0,0016) is statistically significant at 5 % significance level.

## CLUSTER ANALYSIS RESULTS

Several clustering approaches were investigated. Clustering with three-cluster solution as the final partition based on standardized data for all three variables for 34 countries (data for 2013), with the squared Euclidean distance and the Ward Linkage gave the distances between the cluster centroids as it is given in Table 2. The highest is the distance (3,86102) between Cluster 2 with 11 highly developed countries: Austria, Germany, Switzerland, Belgium, France, Ireland, Denmark, Sweden, Netherlands, Finland, United Kingdom, and Cluster 3 with four countries: Spain, Greece, R. Macedonia and Bosnia and Herzegovina.

The three-cluster solution gave the dendrogram given in Figure 6, where four countries, Greece, Spain, R. Macedonia and Bosnia and Herzegovina gathered in separate Cluster 3.

Applying the same clustering method for the same variables and the same 34 countries, even more illustrative is the four-cluster, which gave the distances between cluster centroids as it is given in Table 3. Very high is the distance (3,60798) between Cluster 1 centroid with 6 countries: Albania, Bulgaria, Montenegro, Serbia, Romania, Turkey, and Cluster 2 centroid with 13 countries: Cyprus, Portugal, Croatia, Italy, Poland, Czech R., Malta, Estonia, Slovakia, Hungary, Latvia, Lithuania, Slovenia. However, the highest is the distance (3,86102) between centroids of Cluster 2 and Cluster 3, where there are Bosnia and Herzegovina, R. Macedonia, Spain, and Greece.

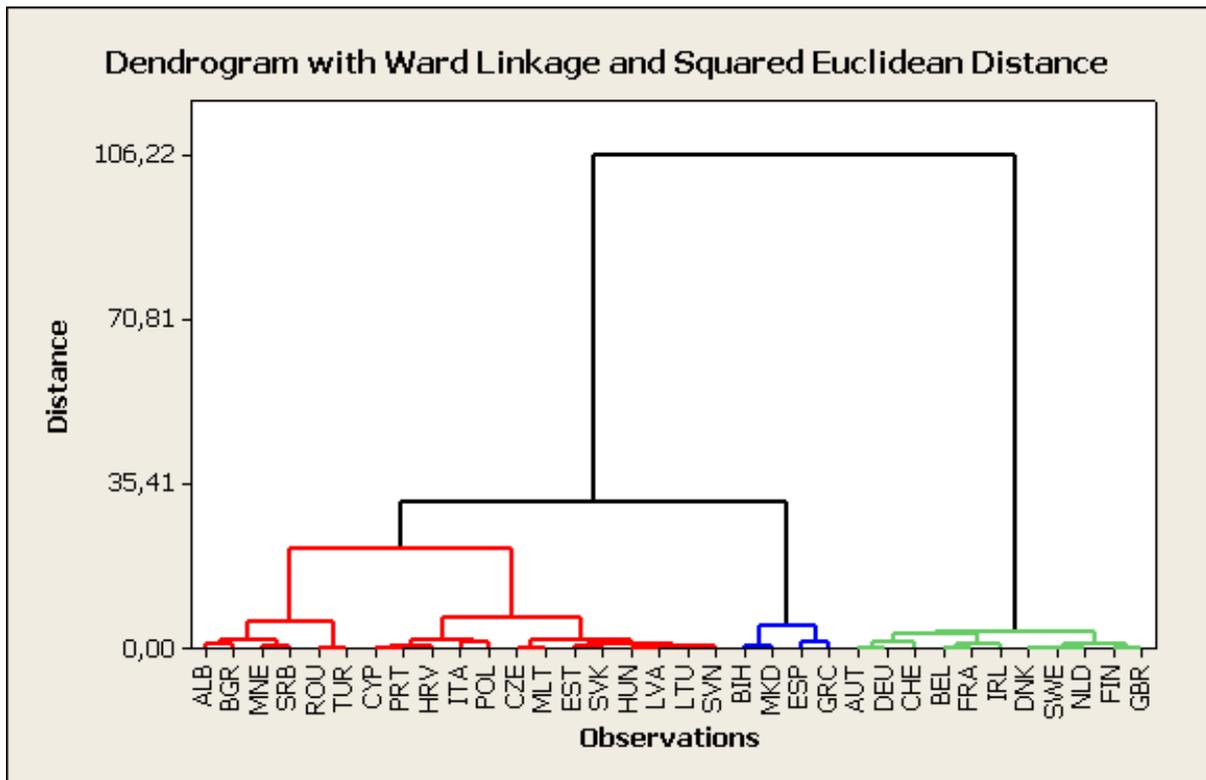
The corresponding dendrogram for the four-cluster solution is given in Figure 7. Cluster of 19 members from Figure 6 is split into two clusters, one with 13 and the other with 6 members.

**Table 2.** The three-cluster solution: Distances between Cluster Centroids.

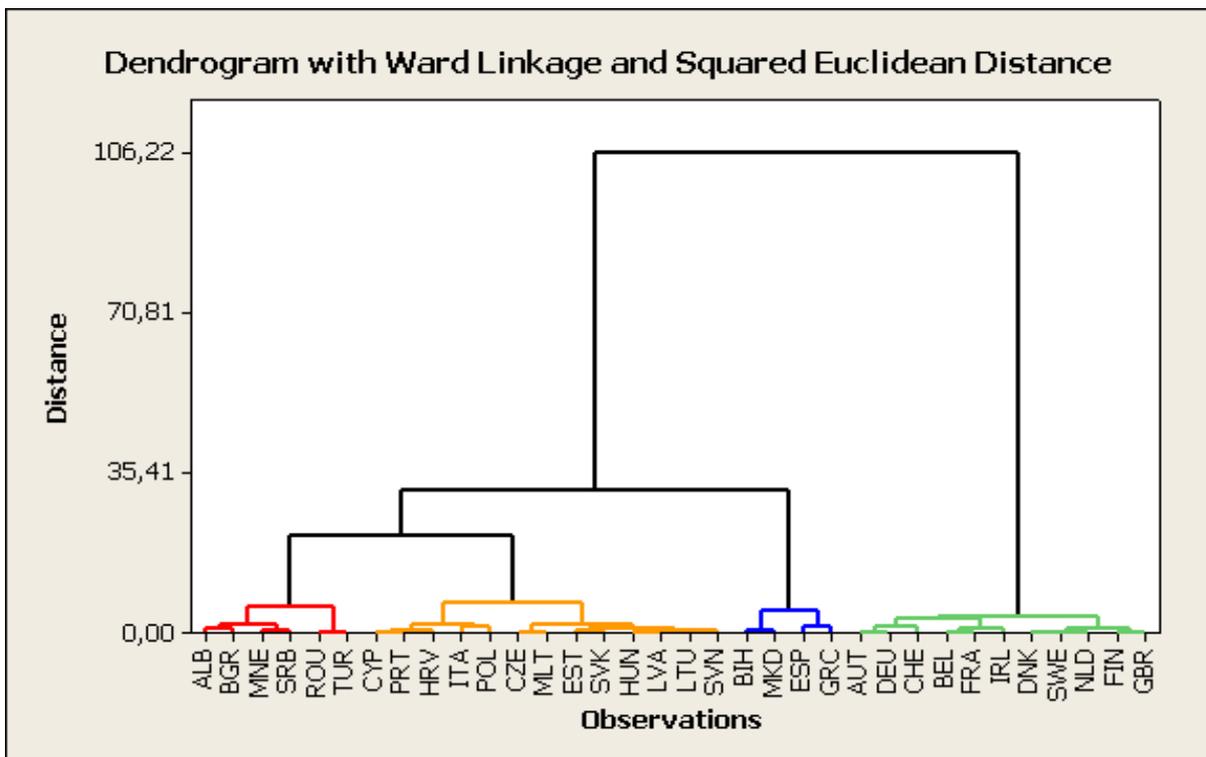
	Cluster1	Cluster2	Cluster3
Cluster1	0,00000	2,51640	2,18576
Cluster2	2,51640	0,00000	3,86102
Cluster3	2,18576	3,86102	0,00000

**Table 3.** The four-cluster solution: Distances between Cluster Centroids.

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,00000			
Cluster2	<b>3,60798</b>	0,00000		
Cluster3	2,11977	<b>3,86102</b>	0,00000	
Cluster4	1,60752	2,01563	2,39264	0,00000



**Figure 6.** The three-cluster solution: The dendrogram based on three standardised variables for 34 countries for 2013. Source: authors' creation and [16].



**Figure 7.** The four-cluster solution: Dendrogram for three standardised variables for 34 countries for 2013: four-cluster solution. Source: authors' creation and [16].

The Table 4 indicates that the SEE countries are not homogeneous while they are scattered over three clusters.

The detailed list of countries comprising each of the clusters shown in Figure 7 is given in Table 4.

**Table 4.** The four-cluster solution: Clusters of countries, three standardised variables for 34 countries for 2013\*. The SEE countries are bolded.

Cluster	No. of countries		Countries
1st	6	<b>SEE and WB countries only</b>	<b>Albania*, Bulgaria, Montenegro*, Serbia*, Romania, Turkey</b>
2nd	13	SEE countries plus post-communist Central and North European plus Mediterranean countries	Cyprus, Portugal, <b>Croatia*</b> , Italy, Poland, Czech R., Malta, Estonia, Slovakia, Hungary, Latvia, Lithuania, Slovenia
3rd	4	<b>SEE and WB countries and Spain</b>	<b>Bosnia and Herzegovina*, R. Macedonia*, Spain, Greece</b>
4th	11	Developed countries only	Austria, Germany, Switzerland, Belgium, France, Ireland, Denmark, Sweden, Netherlands, Finland, United Kingdom

\*denotes the WB countries

## PROFILES OF COUNTRIES

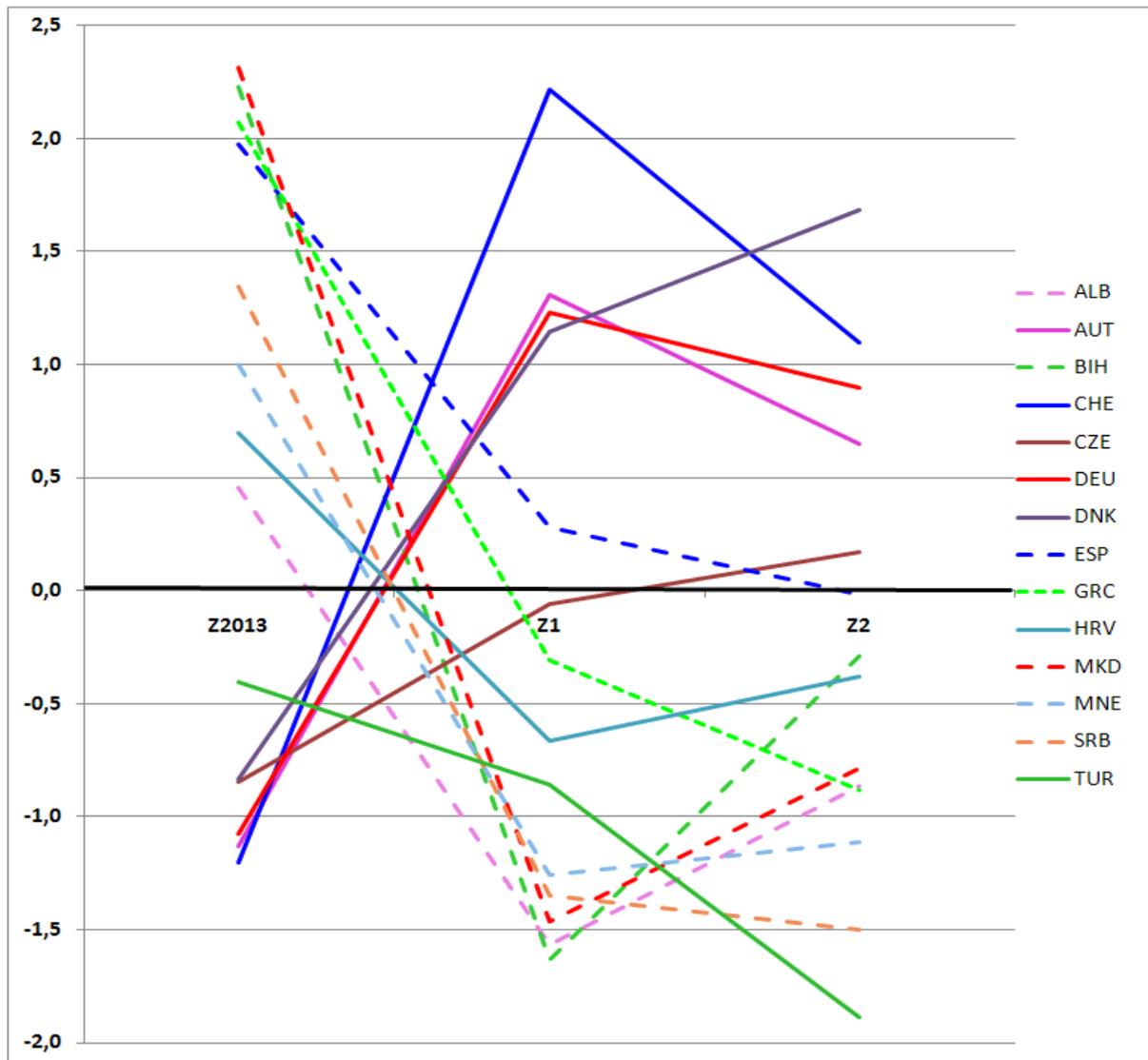
The profile chart is created using calculations of averages and standard deviations for standardised values of variables for 34 countries, but it is shown in Figure 8 only for selected 14 European countries: 7 the most developed and 7 the less developed regarding the considered variables. It indicates that for the variable  $Y_{UemRate}$  is the highest for the countries with the lowest values of the variable  $X_{GDPpc}$ , and vice versa, the highest  $X_{GDPpc}$  values go with the lowest values of the variable  $Y_{UemRate}$ . Regarding  $X_{GDPpc}$  and  $X_{IntUse}$ , four countries are over-averaged and these are Denmark, Switzerland, Germany and Austria, all having very low unemployment rates. It is interesting to notice that Turkey is under-averaged for all three variables, which is good regarding the unemployment rate. At the same time, Spain is over-averaged for all three considered variables, with the variable  $X_{IntUse}$  touching the average line 0,0. Czech Republic is the closest to the average of all three variables. Seven countries, Albania, Bosnia and Herzegovina, Croatia, Greece, R. Macedonia, Montenegro and Serbia, are all with very high unemployment rates and with low GDP per capita and with low level of Internet penetration rates. These countries are all the WB countries, with the exception of Greece.

## CONCLUSIONS

After exploration of 37 European countries' data, only 34 of them remain in the study, since serious outliers for GDP per capita for three countries in almost all the years from 1991 to 2013 appeared.

The World Bank data for 34 countries for 2013 were analysed using multivariate analysis, such as regression modelling and clustering. Multiple regression modelling, firstly, with the original data, and afterwards, with their logarithms is developed in discovering if the unemployment rate would be influenced by the GDP per capita in PPP in current international \$ ( $X_{GDPpc}$ ) and by the Internet penetration rate, defined as the percentage of Internet users per 100 people ( $X_{IntUse}$ ). Since the considered regression models had serious violations of model assumptions, they are not acceptable. Therefore, two simple linear regression models were developed using logarithmically transformed data for all variables, which appeared to give statistically significant models.

The research proves the simple linear regression Model 1 showing the negative correlation with the regressor  $\ln X_{GDPpc}$  and the main variable under study  $\ln Y_{UemRate}$ , with the statistically



**Figure 8.** The profile chart for standardised values: Z2013 ( $Y_{UemRate}$ ), Z1 ( $X_{GDPpc}$ ) and Z2 ( $X_{IntUse}$ ), based on 34 countries averages and standard deviations for the year 2013. Source: authors' creation and [16].

significant estimated regression coefficient of  $\hat{\beta}_1 = -0,75\%$ , explaining 46 % of the total sum of squares, and the regression coefficient of variation is 16 %. Therefore, the estimated model might be considered as moderately representative one. The second simple linear regression Model 2 shows the negative correlation of  $\ln X_{IntUse}$  on  $\ln Y_{UemRate}$  with the estimated regression coefficient of  $\hat{\beta}_1 = -1,38\%$ , explaining 27 % of the total sum of squares, and the regression coefficient of variation is 18 %, being less representative then the first model. All the regression models' assumptions are filled for both models.

Using the cluster analysis based on the Ward linkage and the squared Euclidean distances resulted with the four-cluster solution, showing that the SEE countries are heterogeneous being the members of three clusters. Only Bosnia and Herzegovina, R. Macedonia, Spain and Greece seem to create a very compact cluster, all with very high unemployment rates. In the same time, the most developed European countries included into this research are apart, forming a "compact" cluster of their own, too.

Profile chart for selected seven the less developed and seven the most developed European countries gave proof that European countries differ between each other a lot, but some

clusters might be recognised in this graph, too. Those less developed countries have some similarities, and those the most developed are similar, too. Profile chart indicates that for the variable  $Y_{UemRate}$  is the highest for the countries with the lowest values of the variable  $X_{GDPpc}$ , and the highest  $X_{GDPpc}$  values go with the lowest values of the variable  $Y_{UemRate}$ . Five countries are over-averaged regarding  $X_{GDPpc}$  and  $X_{IntUse}$ , and these are Denmark, Switzerland, Germany and Austria and they all have very low unemployment rates. It is interesting to notice that Turkey is under-averaged for all three variables, which is good for unemployment rate. In the same time, Spain is over-averaged for all three variables, with the variable  $X_{IntUse}$  touching the average line 0,0. Czech Republic is the closest to the average of all three variables. Seven countries: Albania, Bosnia and Herzegovina, Croatia, Greece, R. Macedonia, Montenegro and Serbia are all with very high unemployment rates and with low levels of both GDP per capita and Internet penetration rate.

## ACKNOWLEDGMENT

This work has been supported partially by Croatian Science Foundation under the project STRENGTHS (project no. 9402).

## REFERENCES

- [1] World Bank: *International Comparison Program database*.  
<http://data.worldbank.org>, accessed 15<sup>th</sup> February 2015,
- [2] European Commission: *European Economic Forecast: Spring 2014*.  
[http://ec.europa.eu/economy\\_finance/publications/european\\_economy/2014/pdf/ee3\\_en.pdf](http://ec.europa.eu/economy_finance/publications/european_economy/2014/pdf/ee3_en.pdf),  
accessed 15<sup>th</sup> February 2015,
- [3] Gomez-Salvador, R. and Leiner-Killinger, N.: *An Analysis of Youth Unemployment in the Euro Area*.  
European Central Bank Occasional Paper Series **89**, p.45, 2008,  
<http://www.ecb.europa.eu/pub/pdf/scpops/ecbocp89.pdf>, accessed 19<sup>th</sup> February 2015,
- [4] Kreishan, F.M.: *Economic Growth and Unemployment: An Empirical Analysis*.  
Journal of Social Sciences **7**(2), 228-231, 2011,  
<http://dx.doi.org/10.3844/jssp.2011.228.231>,
- [5] Özel, H.A.; Sezgin, F.H. and Topkaya, Ö.: *Investigation of Economic Growth and Unemployment Relationship for G7 Countries Using Panel Regression Analysis*.  
International Journal of Business and Social Science **4**(6), 163-171, 2013,
- [6] Dovern, J. and Weisser, J.: *Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7*.  
International Journal of Forecasting **27**(2), 452-465, 2011,  
<http://dx.doi.org/10.1016/j.ijforecast.2010.05.016>,
- [7] Bratu, S.M.: *Predicting Macroeconomic Indicators in the Czech Republic Using Econometric Models and Exponential Smoothing Techniques*.  
South East European Journal of Economics and Business **7**(2), 89-99, 2012,
- [8] Bratu, S.M.: *The Accuracy and Bias Evaluation of the Unemployment Rate Forecasts: Methods to Improve the Forecasts Accuracy*.  
Annals of the University of Petrosani – Economics **12**(4), 17-32, 2012,
- [9] Bratu, S.M.: *Forecasts for Inflation and Unemployment Rate Based on Models Using Resample Techniques*.  
International Journal of Economic Practices and Theories **3**(2), 103-107, 2013,
- [10] Simionescu, M.: *The Performance of Unemployment Rate Predictions in Romania: Strategies to Improve the Forecasts Accuracy*.  
Review of Economics Perspectives **13**(4), 161-175, 2013,  
<http://dx.doi.org/10.2478/revecp-2013-0007>,

- [11] Voineagu, V.; Pisica, S. and Caragea, N.: *Forecasting Monthly Unemployment by Econometric Smoothing Techniques*.  
Journal of Economic Computation and Economic Cybernetics Studies and Research **46**(3), 255-267, 2012,
- [12] Eurostat: *Unemployment statistics*.  
[http://ec.europa.eu/eurostat/statistics-explained/index.php/Unemployment\\_statistics#Unemployment\\_trends](http://ec.europa.eu/eurostat/statistics-explained/index.php/Unemployment_statistics#Unemployment_trends), accessed 20<sup>th</sup> February 2015,
- [13] Dumičić, K.: *Developing Forecasting Models for Unemployment Rate by Gender: Cross Countries Comparison*.  
World Statistics Congress ISI'2015, Rio de Janeiro, August 2015, accepted for publication in March 2015,
- [14] Dumičić, K.; Čeh Časni, A. and Žmuk, B.: *Forecasting Unemployment Rate in Selected European Countries Using Smoothing Methods*.  
World Academy of Science, Engineering and Technology: International Journal of Social, Education, Economics and Management Engineering **9**(4), 867-872, 2015,
- [15] International Labour Organization: *Guidelines*.  
<http://www.ilo.org/Search4/search.do?searchLanguage=en&searchWhat=unemployment+person>, accessed 25<sup>th</sup> February 2015,
- [16] Europedia: *Nicholas Moussis*.  
<http://www.europedia.moussis.eu/author>, accessed 6<sup>th</sup> February 2015.

# DISTANCE FACTORS AND CROATIAN EXPORT OBSTACLES IN THE EU15: CAGE APPROACH

Helena Miloloža\*

Faculty of Economics and Business University of Zagreb  
Zagreb, Croatia

DOI: 10.7906/indecs.13.3.8  
Regular article

*Received:* 28 May 2015.  
*Accepted:* 23 June 2015.

## ABSTRACT

For long-term sustainability of any enterprise, it is required to observe each market separately and systematically modify business approach to prevailing conditions in each. However, adaptation to each market can be a complex affair that requires certain financial costs but ignoring distinguishing criteria leads to reduction in revenues and undermines long-term profitability. Time saving and cost-effectiveness can be achieved by application of The CAGE Distance Framework – a model that analysis all diversities between two or more countries.

Croatia did not yet utilize the benefits that European Union membership provides nor did significantly improved its international business. It is obvious that Croatia faces difficulties in the strategic approach to prospective foreign markets. Therefore, emphasis of this article is on analysis and identification of the distinguishing factors between Croatia and the EU15 in order to facilitate and enhance cooperation among them in the future. Throughout the analysis of Croatian export from the beginning of the economic crisis, this article gives an overview of international collaboration among Croatia and the EU15 and comparative analysis by using The CAGE Distance Framework. This analysis has confirmed significant fluctuations, unexploited potential and problems of the Croatian economy in all dimensions of The CAGE model.

## KEY WORDS

global market, CAGE distance framework, Croatia, The EU15, export

## CLASSIFICATION

JEL: F02

\*Corresponding author, *η*: hmiloloza@efzg.hr; +385 1 238 3465;  
Faculty of Economics and Business, Trg J.F. Kennedyja 6, HR – 10 000 Zagreb, Croatia

## **INTRODUCTION**

Export represents a cost-effective way for rapid penetration into new markets and is of a great importance for the development of post-transition economies such as Croatian economy. In terms of internationalization and globalization, it is impossible to maintain and enhance the national economy without an expansive economic strategy. Only export competitiveness creates sustainable competitive advantage and economic growth.

However, the decision about expanding to foreign market is a complex process that depends on many different factors. As a rule, the decision about export on certain market is easier if there is a greater similarity between countries that participate in the international exchange of goods and services. On the other hand, the company will, in general, have more difficulties to export to countries that have different features. The CAGE Distance Framework presents a systematic model that defines all differences between countries through four frames – cultural, administrative, geographic and economic distance. Each of these distances includes various factors. Certain factors are easily recognizable, while others are hidden. The CAGE Distance Framework interprets them as the baseline for defining an export strategy [1].

International exchange of Croatia has always had a negative trend. Croatia cannot overcome the inherent weaknesses and challenges of the export sector, which has been marked by: unsatisfactory technological level, low added value to the products and services, lack of economic and political support and consistent focus on a small number of close markets. However, although only 13 % of companies in Croatia are exporting, they also employ half of the employees in all companies and generate 65 % of revenues [2].

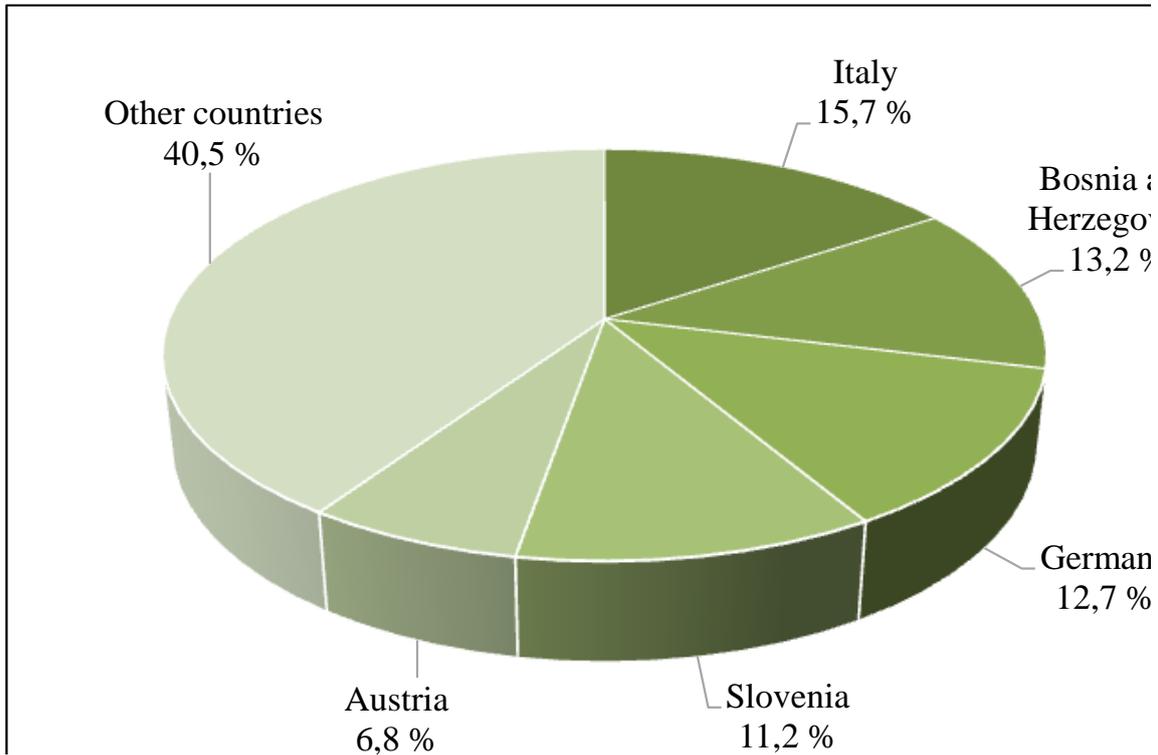
In the modern economy there is a great and decisive importance of trade agreements. Thus, access to the global market becomes easier than ever before. Croatian accession to the European Union has shown its positive effects in the first year of membership. It has been recorded a relatively high growth rates in 2014. The increase in export did not apply only by the increase in export of domestic products but is partly a result of Croatia's mediation in trade between the European Union and third countries. But even though development has been recorded, there is a trend of export retardation, which lasted throughout 2014, and which reflects in the first published data for January 2015 [3]. Therefore, it is obvious that Croatian membership in the European Union has not been utilized and competitiveness of Croatia in the European Union market is still questionable.

The aims of this article are: (i) to highlight the importance of the EU15<sup>1</sup> in the export of Croatia, and (ii) to present The CAGE Distance framework and make a comparison of each member of the EU15 with Croatia. Through a comparative analysis of differences between Croatia and the EU15, it would possible to improve the chances for successful performance of Croatian companies on these profitable unique foreign markets.

## **FOREIGN TRADE BETWEEN CROATIA AND EU15**

Croatia's extroversion towards the European Union is undeniable. The European Union represents the most important foreign trade partner for Croatia. Every year Croatia places around 60 % of its goods and services intended for export on the European Union market. In export, the cooperation with Italy, Germany and Austria prevails, which are, along with Bosnia and Herzegovina and Slovenia, the top five export markets of Croatia (Figure 1).

Croatian economy has established good relations with the EU15 by exchanging various goods and services which can be noticed from Table 1. Croatian export in many EU15 countries fluctuated, but generally tends to increase in the last five years. However, the total value of export to the EU15 has been stabilized at around 5,5 billion US dollars.



**Figure 1.** The main export partners in the foreign trade of the Croatia in 2013 [5]

**Table 1.** Export of Croatia in the EU15 in the period from 2010 to 2014 (thousands of USD) [6].

Country	2010	2011	2012	2013	2014
Austria	624 280	762 140	808 068	797 360	842 120
Belgium	91 011	122 012	136 850	115 589	184 089
Denmark	30 731	34 389	41 564	40 688	44 782
Finland	16 829	28 906	30 399	23 813	23 655
France	161 906	380 752	182 955	219 201	306 537
Greece	88 605	48 742	92 825	118 687	154 422
Ireland	19 291	8 615	9 968	13 159	17 905
Italy	2 209 839	2 111 581	1 892 181	1 851 796	1 916 140
Luxembourg	314 353	331 099	108 663	103 686	22 481
Netherlands	169 241	179 346	207 762	183 942	218 124
Germany	1 218 670	1 349 254	1 260 669	1 496 969	1 538 411
Portugal	8 939	15 061	7 177	14 070	55 683
Spain	79 053	73 066	53 679	73 728	179 042
Sweden	60 600	63 886	59 807	76 866	100 207
UK	175 092	199 745	193 622	273 161	206 053
<b>TOTAL</b>	<b>5 268 440</b>	<b>5 708 593</b>	<b>5 086 187</b>	<b>5 402 714</b>	<b>5 809 650</b>

The main Croatian export products and their share in the total export of the Croatia to each country of the EU15 is shown in Table 2. In export, food, industrial and pharmaceutical products dominate. From industrial products, in export to the EU15 the most important are electrical appliances, petroleum products, steel, iron and cement. From agricultural products, except fertilizers, special interest has been expressed for food. There is also a great proportion of products with high added value (electrical transformers and generators) as well as pharmaceuticals.

**Table 2.** Mayor Croatian export products and their share in the total export of Croatia to each country of the EU15 [7].

Country	The most important export products in 2013	Percentage, %
Austria	Electrical resistors, footwear, insulated wire, roller bearings, diodes, transistors, and similar semiconductor devices	27,7
Belgium	Antibiotics, boxes for metal foundry, leaves for fuming, television, furniture	35,9
Denmark	Products made of aluminium, iron and steel, electrical appliances, wood	42,7
Finland	Steam vapour, electrical transformers, insulated wire, motor vehicles and associated parts	74,7
France	Motor parts, shoes, machines, articles of iron or steel boxes for metal foundry	33,5
Greece	Petroleum oils, fertilizer, sugar, soy	88,9
Italy	Petroleum oils and gases, socks, insulated wire, iron products	23,1
Netherlands	Drugs, electrical transformers, aircraft, iron and steel structures, products for the transport and packing of goods	51,5
Germany	Seats, insulated wires, appliances for automatic regulation, footwear, aluminium products	23,1
Portugal	Fertilizer, leather, pharmaceuticals, aluminium products, central heating boilers	70,7
Spain	Seats, medicines, fertilizers and other chemicals, electric transformers	52,2
Sweden	Prefabricated buildings, electrical transformers, boilers, cement, food and beverages	52,7
UK	Boilers, petroleum oils, insulated wire, soap, food and drink	50,6

The export of services is dominated by tourism which is the dominant export product of Croatia. Namely, tourism is considered as an invisible export because tourists have to go to a different country for their experience where they spend their money. In that case, tourism, as an export industry, sells a significant share of its goods or services outside of the country, thus bringing new money into the local economy. Tourists from the EU15 achieved a total of 12,4 million arrivals and 64,8 million overnight stays in 2013. Their share in the total number of arrivals was 46,9 % and in a total of 51,5 % of total overnight stays in Croatia in 2013 [8].

In a review of the exchange of capital assets, there have been no major Croatian capital investments in the EU15. However, the EU15 countries have invested significant resources in Croatia, which has made a great contribution on development and facilitation of Croatian export. According to the Croatian National Bank, Austria has the status of the most important investor in Croatia [9]. In the period from 1993 to the end of 2013, Austria has invested 7,1 billion euros in Croatia. At the second place, with 12,4 % of the total direct investment in Croatia, is Netherland (3,32 billion EUR), and on the third place is Germany (3,21 billion EUR).

Among the major investors in Croatia emphasizes the Italy with a total of 5,1 % of total foreign investments (1,41 billion EUR) and France with 4,2 % of total foreign direct investments (1,39 billion EUR). The list of top 15 investor also includes Belgium (591,6 mil. EUR), Sweden (529,0 mil. EUR) and United Kingdom (430,4 mil. EUR). On the other hand, Finland, Greece, Ireland and Luxembourg have not made significant investments in Croatia.

According to the research of the Croatian Chamber of Commerce, each of the EU15 countries expressed interest for improving the collaboration with Croatia [10]. The EU15 are highly

developed countries and their interests are technologically appointed. They are interested in energy, IT, telecommunications and high technologies. Opportunities for cooperation exist in the field of metal processing, electrical and electronic industries, and textile and footwear industry. Greenfield investments, innovation, education and the displacement of production and cooperation between different regions of the two countries in the sectors of common interest (agriculture, wood processing, tourism, food industry, etc.) are also areas in which the EU15 sees Croatia as a potential business partner. There is space for improvements in the automotive and aircraft industry as well as in the military equipment production. However, Croatia has the greatest assignment to improve the quality of the industrial sector in terms of quantities and high technological level of production in order to be able to meet the demand by the EU15 countries.

## **CAGE DISTANCE FRAMEWORK**

Differences in business environment are sometimes hard to identify. Diversities often act as a barrier to export activities. Potential international market is established on the basis of unique criteria that each company sets for itself. The attractive market for a particular company can be based on the price of raw materials, transport costs, accessibility options for the selected markets, habits and preferences of consumers, the economic power of the customers or countries and a number of other conditions under which the company makes a decision about participating in certain foreign markets. Of course, any type of international expansion strategy must be supported by specific assets and capabilities that the company possesses [11].

Understanding the differences and their control is greatly facilitated if there is a universal theoretical model in which any company, regardless of their individual characteristics, can incorporate. Since the beginning of the 20<sup>th</sup> century economists were trying to define models that will facilitate the understanding of the international exchange and the existence of comparative advantages of a country or company over another.

Pankaj Ghemawat, a professor at the University in Spain, defined the most comprehensive framework for examining the function and impact of the various diversities on internationalization. The CAGE Distance Framework is the model that provides detailed, comprehensive and non-discriminatory access to perspectives of different kind of distances that can occur between countries on the basis of their own limitations and mutual differences. The CAGE Distance Framework identifies Cultural, Administrative, Geographic and Economic distances between countries that companies should address when crafting international strategies. The most distinctive feature of The CAGE Distance Framework is that it encompasses the bilateral attributes of country pairs as well as the unilateral attributes of individual countries. Table 3 shows a detailed review of each category within The CAGE Distance Framework that includes bilateral and unilateral factors. Bilateral factors are attributes of country pairs that include characteristics and differences which occur by comparing the two countries. Unilateral factors or unilateral attributes of individual countries are related to their self-analysis which helps in identification of personal characteristics that can act as significant aggravating sentence while doing international business.

*Cultural distance* has the greatest impact on consumer preferences as they explain the initial modes of interaction between people, companies and institutions. Different religious belief, racial differences, social patterns of behavior and language barriers can significantly hinder business cooperation between companies [13].

*Administrative distance* is related to differences in the bureaucracy, its work and political structures that are prevailing in two countries. The historical and political background of the country is often crucial in achieving the international cooperation. Governments are the ones

**Table 3.** Four dimensions of The CAGE Distance Framework [12].

<b>Dimension</b>	<b>Country pairs (bilateral)</b>	<b>Countries (unilateral/multilateral)</b>
<b>Cultural Distance</b>	Different languages Different ethnicities Different religions Different values, norms, and dispositions	Insularity Traditionalism
<b>Administrative Distance</b>	Lack of colonial ties Lack of shared regional trading bloc Lack of common currency Political hostility	Nonmarket/closed economy Lack of membership in international organizations Weak institutions, corruption
<b>Geographic Distance</b>	Physical distance Lack of land border Differences in time zones Differences in climates	Landlockedness Lack of internal navigability Geographic size Geographic remoteness Weak transportation or communication links
<b>Economic Distance</b>	Rich/poor differences Other differences in cost or quality of natural, financial and human resources, infrastructure, information or knowledge	Economic size Low per capita income

that can create administrative and political obstacle to other countries by induction of unilateral measures (tariffs, trade quotas, restrictions on foreign direct investment, subsidies to domestic producers or regulating procurement) to foreign competitors. Also, weak institutional infrastructure of the country can be used for the amortization of cross-border activities. On the other hand, companies have an aversion of doing business in the countries that are known for corruption or political conflicts. For these reasons, international integrations and organizations have an important role in the international trade. The European Union is a leading example of a conscious attempt to remove administrative and political differences and facilitate cooperation between member states [14].

*Geographic distance*, except the pith geographical distance, also includes topographic features, the size of the country, the average distance of some cities to national borders and access to waterways. Obviously, the geographic distance affects the cost of transportation, as well as the communication and information barriers. Consequently, its presence may be a dumping measure on trade and investment flows [15]. Therefore, for the companies engaged in the production and trade of solid products or companies whose business requires a high degree of coordination the geographical segment is of great importance.

*Economic distance* is related to differences in economic conditions between two countries. It includes fundamental differences related to income, wealth distribution and the relative purchasing power in certain international markets. Wealth or income of the consumer is the most important economic attribute that creates differences between countries. Moreover, it has a significant impact on the possibility of achieving business cooperation and the level of trade. Since there is a positive correlation between the GDP per capita and trade flows, economically stronger and more stable countries achieve noteworthy international exchange [16]. Also, companies that rely on standardization, experience and economy of scale, reveal their interest to similar economic profile markets because it is easier to implement existing business strategy and well-known patterns of action rather than to implement a completely new strategy on unknown foreign market [17].

Each of the elements of The CAGE Distance Framework “hits” economic relations in a different way and in different proportions depending on the type of industry in which the company operates. However, the basic principle is that: the greater the distance between the countries is present, the greater is the uncertainty of the business.

## **ANALYSIS OF THE CROATIA AND THE EU15 BY USING THE CAGE DISTANCE FRAMEWORK**

The CAGE Distance Framework can be used for better understanding of trade, capital, information and human behaviour patterns. Applying The CAGE Distance Framework, and guided by their own priorities (costs of initial resources, facilitated/difficult access to foreign market and/or consumers, and many other key decision-making criteria), companies can easily study the parameters of the distance within a single model, identify and recognize attractive markets and more clearly see the possibilities and risks of international business.

In the process of internationalization, managers are forced to make decisions about foreign markets. The decision they make, on behalf of the company, is one of the most risky and most crucial, because a single mistake in the selection can have long-term negative effect on business performance. The CAGE Distance Framework can help managers in defining export strategy and to facilitate the export decision by detecting differences between their own company and the market of interest. It helps in the assessment of the effects of distances that act as barriers to the export of various industries in different markets and facilitate acquiring a competitive advantage and provides the possibility of equal competition with domestic firms and enables implementation of quantitative analysis.

Below are given all the characteristics of the EU15 countries placed within The CAGE Distance Framework, analysed by each individual dimensions of the model and compared with the Croatia and its characteristics [18].

### **CULTURAL DISTANCE**

Communication usually represents one of the biggest obstacles to the realization of international business. If two markets are identical to each other by all criteria except the presence of language barrier, business cooperation will be three times more difficult than in the case when there are no communication differences [19]. From Table 4 it is evident that each country of the EU15, including Croatia, has its own language, which, initially, hinders business cooperation. However, the difference in language is easy to establish in advance so companies have enough time to react and reduce the distance in this segment. On the other hand, there are not many religious differences between the EU15 and Croatia because in all countries dominates Christianity.

In a review of migration processes it is visible that the inhabitants of the EU15 have not recorded significant immigrations. Croatia does not represent a productive market for achieving economic prosperity. However, the dominance of emigration movement of Croats has been recorded. The most significant are the emigrations to Germany but also migrations to Austria, France and Italy stand out.

In fact, some cultural distances are much harder to notice, reveal and understand. Social behaviour patterns and adopted system of values are complex cultural characteristics that are invisible and difficult to detect. It is impossible to measure and statistically evaluate them, so it is necessary to establish more intimate business relationship in order to determine these differences. But the starting observatory point in the form of religion can serve as a framework which defines beliefs, patterns and acceptable norms of behaviour and can support the conclusion of small differences between Croatia and countries of the EU15.

However, there are some statistical indicators – Indicators of governance and development – that provide additional measurement of cultural distance between countries, facilitate understanding of the complex cultural factors and help spotting present diversity [20].

*The Human Development Index* (HDI index) measures the average achievements in a country in three basic dimensions: a long and healthy life, access to knowledge and a decent standard of living. The index was created to emphasise that people and their capabilities should be the ultimate criteria for assessing the development of a country, not economic growth alone. Classifications are not absolute, but rather relative to the amount of countries scored each year. The index is classified into four quartiles: (i) Very high, (ii) High, (iii) Medium and (iv) Low. All countries in Table 5., including Croatia, are classified in the highest category according to the HDI index. They are situated in a group of countries that have very high human development.

*Voice and accountability* captures perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media. Point estimates range from about -2,5 to 2,5. Higher values correspond to better governance outcomes. Croatia, according to the Voice and accountability index, lags behind all EU15 countries and is located at the very bottom of the scale. Finland dominates in the freedom of expression and independent decision-making process, while in all other EU15 countries, except Italy and Greece, there is less autonomy of the population. Yet all the countries are ranked between 1.1 and 1.6 which represents a satisfactory value. Only residents of Italy and Greece have a lower degree of freedom; their coefficients of voice and accountability are under 1.

*Rule of law* is dimension that captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Point estimates range from about -2,5 to 2,5. Higher values correspond to better governance outcomes. According to the criteria of the rule of law in the EU15 were recorded minor discrepancies between countries. Sweden, Denmark, Luxemburg, Netherlands, Austria, United Kingdom, Ireland, Germany and Finland have the best established political systems. Not far behind are Belgium, France, Spain and Portugal. In Greece and Italy is by far the highest probability of the occurrence of violence and crime, and Croatia is situated on the last place, trailing the EU15 countries.

## **ADMINISTRATIVE DISTANCE**

Economists argue that the policy of free trade is ideal because it eliminates distortions and increases economic welfare. Therefore, political and administrative similarity contributes to a strong trade relations between two countries. Colonial relations contribute to the growth of trade by more than 900 %, and bilateral trade agreements and treaties, single currency and political union enhance trade by more than 300 % [21].

The biggest distinguishing criterion between Croatia and the EU15 is the currency. Table 5 shows that all countries, except Denmark, Sweden and United Kingdom, use the euro as its currency. For Croatia, it would be desirable that, as a member of the European Union, it becomes a part of the single monetary union. In that case Croatia would avoid exchange losses and exchange rate fluctuations and overcome the problems of growing demand and emission of local currency.

Croatian accession to the European Union should facilitate and enhance cooperation with the EU15. In fact, membership in a particular trade union creates a common market with common policies, rules and regulations. Through membership, a political community (The

**Table 4.** The cultural characteristics of Croatia and the EU15 [18].

Country	Cultural factors								
	Language (spoken by more than 20 %)	Religion, %			Diaspora (1995 – 2004)		Governance and development indicators		
		Christianity	Other	None	From Croatia	In Croatia	Human Development (2011)	Voice & accountability (2010)	Rule of law (2010)
CROATIA	Croatian	93	2	5	–		Very high	<b>0,44</b>	<b>0,19</b>
Austria	German	78	10	12	38 994	1 550	Very high	1,44	1,80
Belgium	Dutch, French, German	100	0	0	32	75	Very high	1,43	1,40
Denmark	Danish	98	2	0	593	39	Very high	1,58	1,88
Finland	Finnish	85	0	15	877	58	Very high	1,97	1,54
France	French	90	6	4	30 591	425	Very high	1,52	1,23
Germany	German	68	32	0	243 614	8 246	Very high	1,35	1,63
Greece	Greek	98	2	0	361	155	Very high	0,90	0,62
Ireland	English, Irish	92	4	4	563	160	Very high	1,34	1,76
Italy	Italian	96	4	0	19 968	1 707	Very high	0,93	0,38
Luxembourg	Luxembourgish	87	13	0	303	7	Very high	1,56	1,82
Netherlands	Dutch	50	2	42	148	131	Very high	1,49	1,81
Portugal	Portuguese	87	0	4	93	322	Very high	1,12	1,04
Spain	Spanish	94	6	0	876	225	Very high	1,14	1,19
Sweden	Swedish	87	13	0	5 726	259	Very high	1,58	1,95
UK	English	72	6	23	7 474	724	Very high	1,31	1,77

**Table 5.** Administrative features of Croatia and the EU15 [18].

Country	Administrative factors				
	Trade Bloc	Currency	Corruption		Legal origin
			Corruption perceptions index (2014)	Control of corruption (2010), %	
<b>CROATIA</b>	<b>EU</b>	<b>HRK</b>	<b>48</b>	<b>59</b>	<b>Germany</b>
Austria	EU	EUR	72	92	Germany
Belgium	EU	EUR	76	90	France
Denmark	EU	DKK	92	100	Scandinavian
Finland	EU	EUR	89	98	Scandinavian
France	EU	EUR	69	89	France
Germany	EU	EUR	79	93	Germany
Greece	EU	EUR	43	56	France
Ireland	EU	EUR	74	93	UK
Italy	EU	EUR	43	57	France
Luxembourg	EU	EUR	82	95	France
Netherlands	EU	EUR	83	98	France
Portugal	EU	EUR	63	81	France
Spain	EU	EUR	60	81	France
Sweden	EU	SEK	87	99	Scandinavian
UK	EU	GBP	78	90	UK

Parliament), free trade zone, single custom union and common market have been established. Croatia is gradually adjusting to the economic policy of the European Union in order to fully overcome all differences. Becoming a member of the European Union, a larger unique market has been opened to Croatia. More severe competition is present but also more opportunities have been given to Croatia in order to achieve economy of scale, as well as opportunities for better use of resources, as well as significant investments.

Corruption is measured with The Corruption Perceptions Index that ranks countries/territories based on how corrupt a country's public sector is perceived to be. It is a composite index, drawing on corruption-related data from expert and business surveys carried out by a variety of independent and reputable institutions. Scores range from 0 (highly corrupt) to 100 (very clean) [22].

The EU15 leads a successful fight against corruption and have a very high value of The Corruption Perceptions Index (60+). In Croatia, corruption is still at high level. Together with Greece and Italy, Croatia has the lowest administrative staff and non-transparent laws and processes. Corruption is further aided by a lack of economic development and income inequality. Corruption of the system prevents the development of fair and efficient market operations. Corrupted Croatian economy does not provide an open and equal access to the market for all potential business partners what impedes international exchange and cooperation. However, Croatia is undertaking various reforms aimed at establishing an equitable administrative and political system. Numerous measures have been taken to eradicate the corruption, but the complexity of the process requires many years of effort, perseverance and resume control.

## GEOGRAPHIC DISTANCE

Based on the geographical features it is impossible to establish uniform rules on the basis of which trade relations will be established. Certainly, they can promote the cooperation but they are not of a crucial matter. Specifically, in geographical terms, Croatia has far more similarities with the EU15 which can be noticed from Table 6.

There are almost no distinguishing features in terms of climatic conditions, as well as time constraints that would prevent the development of mutual trade relations. Also, vicinity to most countries should facilitate international exchange because of reduced time and transportation costs that are conducive for cooperation. However, only Austria, Italy and Germany have represented a high share of Croatian export (35,06 %). Other countries, except United Kingdom, are not in the top ten Croatian export markets which confirm the lack of Croatian competitiveness in the international market and a great lack of demand for Croatian products and services. Unfortunately, Croatia did not utilize the geographical advantage since the EU15 are technologically modern economies with high level of information and communication technologies through which they overcome physical distances to more distant markets that meet their needs.

## ECONOMIC DISTANCE

Observing the economic differences between Croatia and the EU15, economic power of the state is situated in the foreground. It is manifested through certain macroeconomic indicators such as gross domestic product, real annual growth rate of gross domestic product, the rate of unemployment, inflation and international exchanges which is ultimately reflected as the distinction between export and import in the current account balance. Economic characteristics of Croatia and the EU15 are shown in Table 7.

**Table 6.** The geographical features of Croatia and the EU15 [18].

Country	Geographic factors			
	Distance, km	Land area, km <sup>2</sup>	Time zone	Climate zone
<b>CROATIA</b>	–	<b>56 594</b>	<b>1</b>	<b>Temperate</b>
Austria	271	83 871	1	Temperate
Belgium	1024	30 528	1	Temperate
Denmark	1123	43 094	1	Temperate
Finland	1703	338 145	2	Frigid
France	1082	643 427	1	Temperate
Germany	912	357 022	1	Temperate
Greece	1079	131 957	2	Temperate
Ireland	1799	70 273	0	Temperate
Italy	517	301 340	1	Temperate
Luxembourg	847	2 586	1	Temperate
Netherlands	1085	45 543	1	Temperate
Portugal	2201	92 090	1	Temperate
Spain	1702	505 370	1	Temperate
Sweden	1513	450 295	1	Frigid
United Kingdom	1341	243 610	0	Temperate

**Table 7.** Economic characteristics of Croatia and the EU15 [18].

Country	Economic factors			
	GDP per capita (2014),in US\$	Real GDP Growth Rate (2005-2009), %	Human Development Index (2013)	Internet Penetration (2013)
<b>CROATIA</b>	<b>13 598</b>	<b>2</b>	<b>47</b>	<b>66,7</b>
Austria	50 511	1	21	<b>80,6</b>
Belgium	46 930	1	21	82,2
Denmark	59 819	0	10	94,6
Finland	49 151	1	24	91,5
France	42 560	1	20	81,9
Germany	46 251	1	6	84,0
Greece	21 966	2	29	59,9
Ireland	50 478	0	11	78,2
Italy	35 686	-1	26	58,5
Luxembourg	110 665	2	21	93,8
Netherlands	50 793	1	4	84,0
Portugal	21 738	0	41	62,1
Spain	29 882	1	27	71,6
Sweden	60 381	0	12	94,8
UK	40 781	0	14	89,8

The purchasing power of consumers is an important criterion which is reflected in the gross domestic product per capita. From the few empirical papers on the economic distance, it has been concluded that the great economic distance between countries discourages companies to enter foreign markets. Mentioned phenomena has been explained through the perception of economic ability of the customers. Namely, consumers in countries with alike value of gross domestic product per capita, probably, have similar patterns of consumption and are exposed to similar marketing strategies. For example, in most high-income countries the use of credit cards and Internet buying are common, while in low-income countries trade still occurs in the social commerce stores and the services are charged in coins [23].

Studies have shown a positive correlation between the level of GDP per capita and increased trade flows between similar countries [24]. The gravity model also indicates a reduction in trade due to the differences in per capita income [25]. However, purchasing power of consumers is not the only factor that creates the decision to start businesses on certain foreign market. The traditional models (such as H-O model) confirmed that the trade can be motivated and based on the difference in input factors – resources, information and knowledge. In this case, input endowed economy will develop trade with economies with poor resources, like in the case of Croatia and EU15 countries.

Looking at the gross domestic product per capita, a great lag of Croatia is evident. Luxembourg has by far the highest GDP per capita, almost twice the size of Sweden GDP per capita which is on second place in the ranking of the EU15 countries. GDP per capita of other countries in the EU15 is between 30 000 and 50 000 (except Greece), while GDP per capita in Croatia was recorded at only 13 598 US dollars in 2014 [26].

Human Development Index (HDI index), as an indicator of management and development, is explained with cultural diversity, which is shown in absolute values according to four basic categories (very high, high, medium and low). Table 8 contains ranks of each EU15 country in the overall order that includes 187 countries [27]. A higher ranking on the scale implies better human development index. The EU15 are highly ranked in comparison to Croatia which is far behind them.

Internet penetration shows how many residents, in 100 of them, use the Internet. Using the Internet reduces the differences between countries and facilitate access to information and knowledge. It also shows the degree of social development of society, ICT development and the degree of approximation of the overall economy and the internationalization of modern economic trends. Most EU15 countries are forefront of Internet use except Italy, Portugal and Spain which are at par with the Croatia. Yet, Greece is behind Croatia with only 59.9 people with access to the worldwide network [28].

## CONCLUSIONS

With the development of different models of financing and establishing trade agreements access to the global market has become easier than ever. Globalization has forced countries to open their economies because only by internationalization they can maintain on the market. Because of the low intensity of maturation of the domestic market and the limited possibilities of the same, export is of particular importance to businesses without the financial and managerial resources needed for more extensive international operations, including joint ventures and foreign direct investment.

In modern economies, the structure and the amount of export and import is the result of the application of technological advances, organizational superiority and managerial skills, and on that basis, a lower costs with an emphasis on high quality and product differentiation. The products with high added value, as well as the greater coverage of import by export, enable survival in the inherently complex environment and provide long-term growth.

Croatian economy has expressed the need to integrate into the world economy because Croatia is counted as medium resource-scarce country. Therefore, in order to achieve certain economic effects on international markets, Croatian companies have to provide a high quality products and services with reasonable prices especially on mature markets such as the EU15 market. In addition, Croatia must focus on this market which represents a majority share of Croatian export.

However, trade between Croatia and other countries is very unfavourable which impedes economic development of closed Croatian economy. The biggest problem is that Croatia does not based its positive results on structural reforms, but on the reduction of overall economic activity, overall demand and changes in trends in the prices of raw materials and finished products on the world market. Croatia needs to make quality and sustainable economic policy, as the synergy of fiscal, monetary and other policies, which will, consequently, boost export and reduce the trade deficit. This includes structural reforms that yet have not been intensified.

The goal of the article was to compare Croatia and EU15 countries, using the CAGE Comparator [18]. In the comparison of Croatia and EU15, according to The CAGE Distance Framework, significant fluctuations, unexploited potential and problems of the Croatian economy are visible in all four dimensions of The CAGE model.

Cultural differences between the EU15 and the Croatia are present in terms of language, expression freedom, association and action, as well as in the system of political power which is underneath the level of the EU15. These negative differences act as obstacles in cooperation and decrease Croatian export in the EU15 countries.

Croatian accession to the European Union reduced administrative differences significantly. International trade with the member states has been significantly facilitated. Taxes, tariffs and other barriers that created adverse effects on trade have been reduced and/or eliminated. Croatia is gradually adopting homogeneous rules and regulations of the European Union that should establish order and eradicate the highly pervasive corruption in the Croatian administrative and political system.

In geographic terms, Croatia is very close to the EU15. Croatia, however, has not used geographical proximity as its strategic advantage. Unfortunately, in the past, the geographic distance represented an important issue and was an important factor in achieving business cooperation between distant markets. But in modern times, technology and the Internet access have greatly contributed to overcome these barriers and reduce the distance in terms of time needed to transport goods. Also, with digital products and services, the geographical distance has almost been completely eliminated as a restriction for trade between countries. Therefore, despite the proximity and geographically important position, Croatia does not represent an important business partner for the EU15. Croatian economy faces the fragmentation of production and lack of production capacities so it cannot meet the high demands of the EU15. Most of the Croatian export comes from labor-intensive activities, and competitiveness in such a manner cannot be secured or maintained. In Croatia, there is an indispensable need for structural export changes in terms of increasing the export of products based on new technologies.

The most pronounced differences between the EU15 and Croatia are from an economic point of view. In economic sense, Croatia is a developing country with scarce resources. The low level of education, unskilled labor and low quality of products and services, due to lack of technological development, prevents Croatia to become a serious business partner and equal participant in the international market. The EU15 (except Greece in certain segments) are highly developed countries with regulated economic policies and sustainable economic systems. Croatia is not in such a position and it is a question of time when will Croatia be capable of overcoming the present obstacles.

The Croatian economy must internationalize. Croatia needs to establish a proper legislative and legal basis, define and systematically implement structural reforms, stimulate export activities through the production and finally, design high-quality export strategy in order to achieve significant results on international markets by placing competitive products and services on the promising market such as the EU15 market.

## **REMARK**

<sup>1</sup>The EU15 comprised the following 15 countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom.

## **REFERENCES**

- [1] Ghemawat, P.: *Distance still matters*. Harvard Business Review **79**(8), 137-147, 2001,
- [2] Croatian Ministry of Foreign and European Affairs: *About Croatian Exports*. In Croatian. <http://gd.mvep.hr/hr/o-hrvatskom-izvozu>, accessed 27<sup>th</sup> May 2014,
- [3] Croatian Bureau of Statistics: *Released Data – First Results*. <http://www.dzs.hr>, accessed 13<sup>th</sup> April 2015,
- [4] Croatian Bureau of Statistics: *Statistical Yearbook of the Croatia 2014*. [http://www.dzs.hr/Hrv\\_Eng/ljetopis/2014/sljh2014.pdf](http://www.dzs.hr/Hrv_Eng/ljetopis/2014/sljh2014.pdf), accessed 15<sup>th</sup> April 2014,
- [5] Croatian Bureau of Statistics: *Export and import by partner countries*. In: Statistical Yearbook of the Croatia 2014, Croatian Bureau of Statistics, Zagreb, 2014, [http://www.dzs.hr/Hrv\\_Eng/ljetopis/2014/sljh2014.pdf](http://www.dzs.hr/Hrv_Eng/ljetopis/2014/sljh2014.pdf), accessed 15<sup>th</sup> April 2014,

- [6] Croatian Bureau of Statistics: *Foreign trade in goods*.  
<http://www.dzs.hr>, accessed 15<sup>th</sup> April 2014,
- [7] Croatian Chamber of Economy, International relations sector: *Information about foreign markets*.  
<http://www.hgk.hr/sektor-centar/sektor-medjunarodni/informacije-o-inozemnim-trzistima>, accessed 21<sup>st</sup> April 2014,
- [8] Croatian Bureau of Statistics: *Tourist arrivals and nights, by country of residence*.  
In: Statistical Yearbook of the Croatia 2014, Croatian Bureau of Statistics, Zagreb, 2014,  
[http://www.dzs.hr/Hrv\\_Eng/ljetopis/2014/sljh2014.pdf](http://www.dzs.hr/Hrv_Eng/ljetopis/2014/sljh2014.pdf), accessed 15<sup>th</sup> April 2014,
- [9] Croatian National Bank: *The theory and reality of foreign direct investment in the world and in transition countries with special reference to Croatia*. In Croatian.  
<http://www.hnb.hr/publikac/pregledi/p-009.pdf>, accessed 23<sup>rd</sup> April 2014,
- [10] Croatian Chamber of Economy: *Information about foreign markets*.  
<http://www.hgk.hr/sektor-centar/sektor-medjunarodni/informacije-o-inozemnim-trzistima>, accessed 22<sup>nd</sup> April 2015,
- [11] Malhotra, S.; Sivakumar, K. and Zhu, P.: *Distance factors and target market selection: the moderating effect of market potential*.  
International Marketing Review **26**(6), 651-673, 2009,  
<http://dx.doi.org/10.1108/02651330911001332>,
- [12] Ghemawat, P.: *World 3.0: Global Prosperity and How to Achieve it*.  
Harvard Business Review Press, Boston, 2011,
- [13] Griffith, D.A. and Dimitrova, B.V.: *Business and Cultural Aspects of Psychic Distance and Complementarity of Capabilities in Export Relationships*.  
Journal of International Marketing **22**(3), 50-67, 2014,  
<http://dx.doi.org/10.1509/jim.14.0019>,
- [14] Balassa, B.: *The Theory of Economic Integration*.  
Routledge Revivals, Routledge & Oxford, 2013,
- [15] Bernard, A.B.; Grazi, M. and Tomasi, C.: *Intermediaries in International Trade: direct versus indirect modes of export*,  
w17711, National Bureau of Economic Research, 2011,  
<http://www.nber.org/papers/w17711.pdf>,
- [16] Waugh, M.: *International trade and income differences*.  
The American Economic Review **100**(5), 2093-2124, 2010,  
<http://dx.doi.org/10.1257/aer.100.5.2093>,
- [17] Cassey, A.J.; Holland, D.W. and Razack, A.: *Comparing the economic impact of an export shock in two modeling frameworks*.  
Applied Economic Perspectives and Policy **33**(4), 623-638, 2011,  
<http://dx.doi.org/10.1093/aep/ppr033>,
- [18] Ghemawat, P.: *CAGE Comparator*.  
<http://www.ghemawat.com/cage>, accessed 22<sup>nd</sup> April 2015,
- [19] Hutzschenreuter, T.; Voll, J.C. and Verbeke, A.: *The impact of added cultural distance and cultural diversity on international expansion patterns: A Penrosean perspective*.  
Journal of Management Studies **48**(2), 305-329, 2011,
- [20] Transparency International: *Corruption by Country/Territory*.  
<http://www.transparency.org/country>, accessed 22<sup>nd</sup> April 2015,
- [21] Ghemawat, P.: *Arbitrage Strategies*.  
Unpublished Globalization Note Series, University of Navarra, IESE Business School in Barcelona, Barcelona, 2014,
- [22] Alam, M.S.: *A theory of limits on corruption and some applications*.  
Kyklos **48**(3), 419-435, 1995,  
<http://dx.doi.org/10.1111/j.1467-6435.1995.tb02323.x>,

- [23] Mitra, D. and Golder, P.: *Whose culture matters? Near-market knowledge and its impact on foreign market entry timing*,  
Journal of Marketing Research **39**(3), 350-365, 2002,  
<http://dx.doi.org/10.1509/jmkr.39.3.350.19112>,
- [24] Bergstrand, J.H. and Jeffrey, H.: *The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence*.  
The Review of Economics and Statistics **67**(3), 474-481, 1985,  
<http://dx.doi.org/10.2307/1925976>,
- [25] James, E.A. and Wincoop, V.: *Gravity with Gravitas: A Solution to the Border Puzzle*.  
The American Economic Review **93**(1), 170-192, 2003,  
<http://dx.doi.org/10.1257/000282803321455214>,
- [26] The World Bank: *GDP Per Capita (current US\$)*.  
<http://data.worldbank.org/indicator/NY.GDP.PCAP.CD>, accessed 14<sup>th</sup> June 2015,
- [27] United Nations Development Programme: *Human Development Index*.  
<http://hdr.undp.org/en/content/human-development-index-hdi-table>, accessed 14<sup>th</sup> June 2015,
- [28] The World Bank: *Internet users (per 100 people)*.  
<http://data.worldbank.org/indicator/IT.NET.USER.P2>, accessed 14<sup>th</sup> June 2015.

# THE EFFECT OF SHOCKS: AN EMPIRICAL ANALYSIS OF ETHIOPIA

Yilebes Addisu Damtie\*

Department of Disaster Risk Management and Sustainable Development, Gambella University  
Gambella, Ethiopia

DOI: 10.7906/indecs.13.3.9  
Regular article

*Received:* 27 May 2015.  
*Accepted:* 21 July 2015.

## ABSTRACT

Besides striving for the increase of production and development, it is also necessary to reduce the losses created by the shocks. The people of Ethiopia are exposed to the impact of both natural and man-made shocks. Following this, policy makers, governmental and non-governmental organizations need to identify the important shocks and their effect and use as an input. This study was conducted to identify the food insecurity shocks and to estimate their effect based on the conceptual framework developed in Ethiopia, Amhara National Regional State of Libo Kemkem District. Descriptive statistical analysis, multiple regression, binary logistic regression,  $\chi^2$  and independent sample t-test were used as a data analysis technique. The results showed eight shocks affecting households which were weather variability, weed, plant insect and pest infestation, soil fertility problem, animal disease and epidemics, human disease and epidemics, price fluctuation problem and conflict. Weather variability, plant insect and pest infestation, weed, animal disease and epidemics created a mean loss of 3,821.38, 886.06, 508.04 and 1,418.32 Birr, respectively. In addition, human disease and epidemics, price fluctuation problem and conflict affected 68.11%, 88.11% and 14.59% of households, respectively. Among the sample households 28,1 % were not able to meet their food need throughout the year while 71,9 % could. The result of the multiple regression models revealed that weed existence ( $\beta = -0,142, p < 0,05$ ), plant insect and pest infestation ( $\beta = -0,279, p < 0,01$ ) and soil fertility problem ( $\beta = -0,321, p < 0,01$ ) had significant effect on income. Asset was found significantly affected by plant insect and pest infestation ( $\beta = -0,229, p < 0,01$ ), human disease and epidemics ( $\beta = 0,145, p < 0,05$ ), and soil fertility problem ( $\beta = -0,317, p < 0,01$ ) while food production was affected by soil fertility problem ( $\beta = -0,314, p < 0,01$ ). Binary logistic regression model revealed that food availability of the households was highly affected by the asset ( $\text{Exp}(B) = 1,00, p < 0,1$ ), and food production ( $\text{Exp}(B) = 1,379, p < 0,01$ ).

## KEY WORDS

shocks, effect, Ethiopia, food availability, binary logistic, multiple regression

## CLASSIFICATION

JEL: C12, D13, H31

\*Corresponding author, *η*: yileaaddisu@gmail.com; ;  
Gambella University, P.O. Box – 126, Gambella, Ethiopia

## **INTRODUCTION**

A shock is defined as a ‘sudden event that impacts on the vulnerability of a system and its components’. In case of slow onset hazards it is ‘when the event passes its tipping point and becomes an extreme event’ [1]. Shocks are natural, social, economic, and political in nature. They can occur as slow or rapid onset shocks or longer-term stresses or trends and can be idiosyncratic or covariate in nature. Shocks can be transitory, seasonal, or structural, and their frequency, severity and duration can vary widely [2]. Shock includes human health shocks, natural shocks, economic shocks, conflict, and crop and livestock health shocks [3].

Shocks are usually localized and therefore threaten the livelihoods of only parts of the population directly affected through loss of household assets, market access, and income earning opportunities, among others [4]. Shocks can destroy assets directly (in the case of floods, storms, civil conflict, etc.). They can also force people to abandon their home areas and dispose off assets (such as land) prematurely as part of coping strategies [3].

The Horn of Africa is acutely vulnerable to food security crises that arise from complex causes, including swift shocks from the vagaries of climate, particularly exposure to drought and flooding, and slower moving stresses like the complex nexus of rapid population growth, land fragmentation, natural resource degradation, and conflict [5].

Due to recurrent natural and manmade hazards, degradation of natural resources, lack of land and labor fertility and other related reasons many Ethiopian people live with food insecurity problems [6]. Even though there are some variations across regions drought, flood, erosion, frost, crop pests, livestock pests, input access, input price rise, death and illness are the most important shocks in Ethiopia [7].

The study area, Libo Kemkem, is among the food insecure areas of Amhara region. Productive Safety Net Program and other interventions have been practiced for food insecure households. However, they could not build their resilience to protect themselves from shocks and food insecurity is still the major problem of the area.

The identification of shocks and their effect should be a prerequisite and an input for food security building activities. Despite this, it does not gain adequate consideration neither at country level nor at local level. In view of this, study was conducted to identify the shocks in the Libo Kemkem district and to estimate their effect.

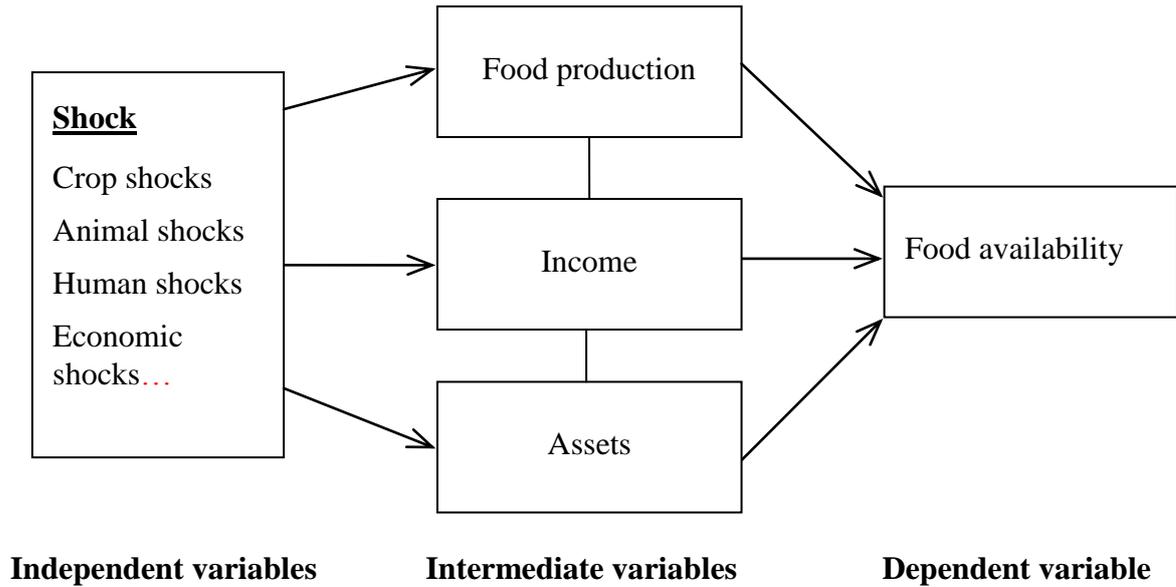
## **CONCEPTUAL FRAMEWORK OF THE STUDY**

Conceptual framework was constructed to show the relationship between variables and to shape the study. The independent variables, crop shocks, animal shocks, human shocks and economic shocks affect food production, income and asset of the household directly and food availability indirectly, Figure 1. Households affected by shocks could have low food production, income and assets possession and this affects household food availability status from production, purchase or shared out negatively. The working hypothesis of this study is stated as: shocks have direct effect on food production, income and assets and indirect effect food availability of households.

## **METHODOLOGY**

### **STUDY AREA**

This study was undertaken in Ethiopia, Amhara National Regional State of Libo Kemkem District. Based on the data from Libo Kemkem District Agriculture and Rural Development



**Figure 1.** Conceptual framework of the study.

Office [8], the population of the district is around 225 499. The district is adjacent to West Belesa in the South, Fogera and Ebenat in the West and East, respectively, and Lake Tana in the West [9]. Addis Zemen is the center of the district and it is found 652 km from the country capital Addis Ababa and 80 km from the regional capital Bahir Dar.

## METHODS OF SAMPLING AND DATA COLLECTION

Data of this study was obtained from field survey. Using two stage random sampling technique 185 households were selected for the study. First, three kebeles were randomly selected as a representative of the district. Then, the sample households were selected proportionally from each kebele. A semi-structured questionnaire was used as a data collection tool. Data collectors randomly selected the households to be interviewed.

## DATA ANALYSIS TECHNIQUE

Data obtained was analyzed through descriptive statistical analysis, multiple regression, binary logistic regression, independent sample t-test and  $\chi^2$ . Descriptive statistics were computed to describe some of the variables in the form of mean, standard deviation and percentage. Independent sample t-test was used for the purpose of comparison of mean differences between food available and food non available households with regard to continuous variables. Likewise  $\chi^2$  was used for dummy variables. The data analysis was done using SPSS version 16.

The multiple regression models were used for this study to analyze the effect of shocks on intermediate variables. Multiple regression is the instrument of choice when the researcher believes that several independent variables interact to predict the value of a dependent variable and when the dependent variable is measured on continuous scale. Thus, three multiple regression models were run to observe the effect of shocks on income, food production and assets.

The general formula for multiple regression models is given as follows

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon, \quad (1)$$

where  $Y_i$  denotes the dependent variable,  $\beta_0$  is the constant term,  $X_1$  to  $X_n$  denote the explanatory variables,  $\beta_1$  to  $\beta_n$  are the coefficients associated with explanatory variables and  $\varepsilon$  is the observable random error term or disturbance.

The multiple linear regression models were estimated using the ENTER method and F-test computed to observe the significance of the models. Diagnostic tests were also carried out to check for multicollinearity of the variables included in the model.

Binary logistic regression was used to estimate the effect of intermediate variables on the dependent variable (food availability). The dependent variable is measured in terms of dummy variables; 0 for food non available and 1 for food available, based on their food availability status throughout the year. The analysis of the logistic regression model shows that, changing an independent variable alters the probability that a given household becomes food available. The equation of logistic regression is

$$Z_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + U_i, \quad (2)$$

where  $\beta_0$  is an intercept,  $\beta_1, \beta_2, \dots, \beta_n$  are slopes of the equation in the model,  $X_1, X_2, \dots, X_n$  are intermediate variables and  $U_i$  is a disturbance term.

## **RESULTS AND DISCUSSION**

### **FOOD AVAILABILITY SITUATION**

Food availability refers to the physical presence of food stocks in desired quantities [10]. The results showed that of the total 185 households, 52 households (28,1 %) were not able to cover their food need by their own throughout the year. The remaining 133 households (71,9 %) had the ability to cover their food need all over the year.

### **THE EXISTENCE AND EFFECTS OF SHOCKS**

The study identified eight shocks and among these weather variability, weed, plant insect and pest infestation and soil fertility problem were crop shocks. Animal disease and epidemics was animal shock and price fluctuation problem was economic shock. In addition, human disease and epidemics and conflict were categorized under human shocks<sup>1</sup>.

The effect of the shocks on food production was measured by comparing the gained production with that of expected production if that shock never happened. Since the households cultivate different type of crops, it is difficult to estimate their crop production loss in kilograms and sum. Thus, the extent of each crop loss converted into its respective price value and summed to gain the total crop loss in Birr<sup>2</sup>.

More than three-fourths (83,78 %) of the sample households were affected by weather variability. The sample households lost a mean of 3 821,38 Birr estimated production damage due to early emergence, delay, excessiveness of rain (flood) or any other form of weather variability.

Plant insect and pest infestation was recorded on 112 (60,54 %) sample households while weed had an effect on 82 (44,32 %) households. Plant insect and pest infestation and weed also created a mean loss of 886,08 Birr and 508,48 Birr, respectively.

In addition, animal disease and epidemics affected 98 (52,97 %) sample households and create a mean loss of 1418,32 Birr.

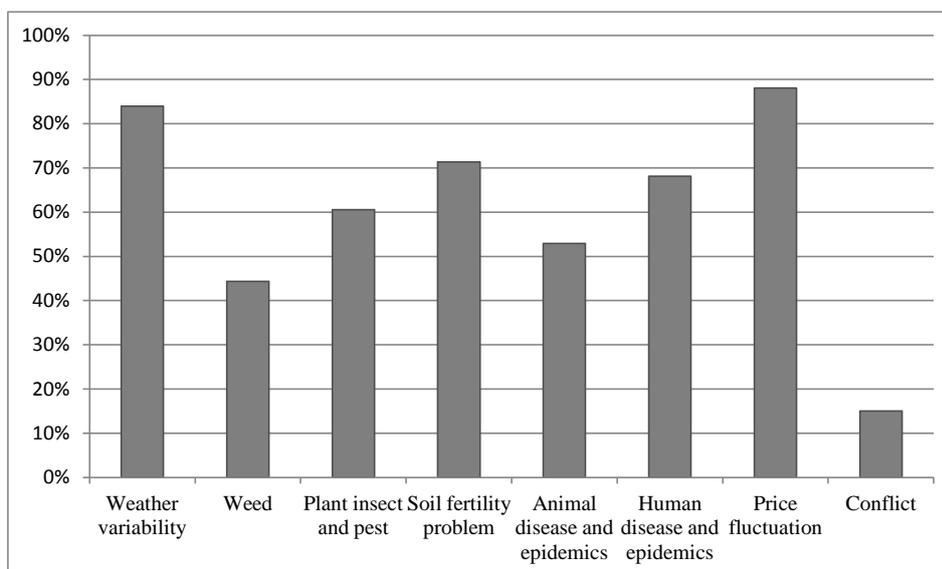
Further, two human shocks affecting households were identified. These shocks, human disease and epidemics and conflict affected 126 (68,11 %), and 27 (14,59 %) sample households, respectively. The sample households had a mean illness score of 0,615. In addition, the economic shock price fluctuation problem had affected 163 (88,11 %) households (Table 1).

Figure 2 shows the existence of the shocks in households. Price fluctuation problem was the primary shock existing in 88,11 % of households. Weather variability and soil fertility problem were the second and third shocks in occurrence, respectively. In contrast, conflict was recorded as the least existing shock occurring only in 14,59 % of sample households.

Table 2 depicts the mean score of the sample households with regard to the effect of the shocks aggregated by food availability. Independent sample T-test and  $\chi^2$  were employed to show the relationship between the food availability and the shocks. The results indicated that only soil fertility problem had statistically significant relationship with food availability at less than 1 % probability level.

**Table 1.** The existence and effect of shocks on sample households ( $N = 185$ ).

No	Type of shock	Existence of the shock				Mean effect of the shock, Birr	Total effect, Birr
		No		Yes			
		No.	%	No.	%		
<b>Crop</b>							
1	Weather variability	30	16,22	155	83,78	3 821,38	70,6955
2	Weed	103	55,68	82	44,32	508,48	94,0680
3	Plant insect and pest infestation	73	39,46	112	60,54	886,08	16,3925
4	Soil fertility problem	53	28,65	132	71,35		
<b>Animal</b>							
5	Animal disease and epidemics	87	47,03	98	52,97	1 418,32	262,389
<b>Human</b>							
6	Human disease and epidemics	59	31,89	126	68,11	0,6149 illness score	113,76 illness score
7	Conflict	158	85,41	27	14,59	0,21	39
<b>Economic</b>							
8	Price fluctuation	22	11,89	163	88,11		



**Figure 2.** The existence of the shocks in sample households in percents ( $N = 185$ ).

**Table 2.** Distribution of sample households by food availability and shocks ( $N = 185$ ).

Type of shock	Variable type	Variable definition	Food non-available ( $N = 52$ )		Food available ( $N = 133$ )		Total sample ( $N = 185$ )		t-value ( $\chi^2$ )
			Mean	SD	Mean	SD	Mean	SD	
Weather variability	Continuous	Crop production loss due to the shock, in Birr	5 833,94	20 583,59	3 034,51	3 452,25	3 821,38	11 294,97	-1,521
Weed	Continuous	Crop production loss due to the shock, in Birr	485,15	826,69	517,59	1092,30	508,48	1022,53	0,193
Plant insect and pest infestation	Continuous	Crop production loss due to the shock, in Birr	745,58	1651,91	941,02	1862,25	886,08	1803,33	0,662
Soil fertility problem	Dummy	0, no soil fertility problem, 1 otherwise							6,225*
Animal disease and epidemics	Continuous	Animal production loss due to the shock, in Birr	1660,23	3 293,13	1323,74	3 572,86	1418,32	3490,93	-0,588
Human disease and epidemics	Continuous	The effect of human disease and epidemics in terms of illness score	0,75	0,74	0,56	0,71	0,615	0,722	-1,571
Price fluctuation problem	Dummy	0, no price fluctuation problem, 1 otherwise							1,218
Conflict	Continuous	Number of conflicts	0,23	0,65	0,20	0,57	0,211	0,593	-0,285

\*significant at probability level 1 %

## FOOD PRODUCTION, INCOME AND ASSET

### Description of and correlation between food production, income and asset

The sample households gained an annual mean income of 5 864,95 Birr. The Food available households earned a mean of 6 869,29 Birr while the food non available households gained 3 296,17 Birr. Besides, the sample households possessed a mean of 11 753,46 Birr estimated assets while 6 093,21 Birr and 13 966,49 Birr were the estimated asset possessions of the non food available and food available households, respectively.

In addition, the households produced food which was sufficient enough for a mean of their 8,99 months food need. The food non available households produced food for a mean of 7,13 months while the food available households produced a mean of 8,99 months food. The independent sample T-test result shows that there was a significant difference in the mean of food non available and food available households at less than 1 % probability level in the income, asset and food production (Table 3).

In order to identify the degree of association between intermediate variables correlation analysis was computed. All of the 3 cells in the correlation matrix were positively and significantly correlated at less than 1 % probability level (2-tailed) (Table 4).

**Table 3.** Distribution of sample households by income, assets and food production ( $N = 185$ ).

Variable	Food non-available ( $N = 52$ )		Food available ( $N = 133$ )		Total sample ( $N = 185$ )		t-value
	Mean	SD	Mean	SD	Mean	SD	
Income, Birr	3 296,17	2 818,24	6 869,29	9 162,55	5 864,95	8 063,63	-2,757*
Assets, Birr	6 093,21	5 866,81	13 966,49	16 310,82	11 753,46	14 594,22	-6,477*
Food production, months	7,13	2,41	9,71	2,45	8,99	2,693	-3,391*

\*significant at probability level 1 %

**Table 4.** Correlation between assets, income and food production ( $N = 185$ ).

Variable	1	2	3
Food production (1)	1		
Income (2)	0,359*	1	
Assets (3)	0,395*	0,616*	1

\*significant at probability level 1 % (2-tailed)

Food production had a positive relation with income of the household ( $r = 0,359$ ,  $p < 0,01$ ) and asset ( $r = 0,395$ ,  $p < 0,01$ ). This is due to the possibility of converting food production in to income. Some of the produced food which is beyond the household need will be taken to the local market and sold. In addition, assets are used as an input for food production activities. Human labour, livestock and farming equipments can be used as an input. Also, fertilizers, pesticides and herbicides are purchased for crop production.

Assets showed a strong positive relationship with income of the household ( $r = 0,616$ ,  $p < 0,01$ ). The possible reason for this is the ability of assets to be used as a generation of income source for the household. Livelihoods of the households depend on their asset possession and these livelihoods create income for the household. Also income could be used as a source of asset. The income could be vested for the creation or development of assets.

### **The effect of shocks on food production, income and assets**

Multiple linear regression models were used to estimate the effect of shocks on intermediate variables. Three multiple linear regression models, for income, asset and food production, were computed (Table 5). These three models showed 22,9 %, 19 % and 12,7 % of the variation, respectively. Multicollinearity tests conducted showed that there was no strong correlation between variables. The result of F-test revealed the significance of all of the models at less than one percent significance level.

Out of eight shock variables included in the income model, three shocks were found to have a significant effect on income of the household. Plant insect and pest infestation ( $\beta = -0,279$ ) and soil fertility problem ( $\beta = -0,321$ ) were found significant at less than one percent probability. In addition, weed existence ( $\beta = -0,142$ ) was found to have a significant effect on income at less than 5 % probability level.

The direction of coefficient of all of these significant variables showed a negative relation with income of the household. With constant condition of other variables, a one standard deviation unit increase in plant insect and pest infestation, soil fertility problem and weed existence resulted in the decrease of household income by 0,279, 0,321 and 0,142 standard deviation units, respectively.

There are some possible reasons for this. First the household incur some cost for the management of these shocks. Insecticide, pesticide, herbicides and fertilizer invested to manage these shocks. Also employed human power could be used for weeding. These are additional costs which has effect on net income of the households. In addition, agricultural land with such shocks creates low income in rent out.

Likewise, the second multiple regression model, asset, come up with three significant shocks. Plant insect and pest infestation ( $\beta = -0,229$ ) and soil fertility problem ( $\beta = -0,317$ ) were found to negatively and significantly affect assets at less than 10 % probability level. This is due to the damage created by the pests and insects on stored agricultural output. Also some pests and insects reduce the quality of agricultural land. In addition, soil fertility problem reduces the quality of agricultural land which is natural asset.

However, human disease and epidemics ( $\beta = 0,145$ ) were found to positively and significantly affect asset at less than 5 % probability level. The result of this shock is in contrast to expected. The possible reason for this is the frequent occurrence of disease and epidemic on the non productive segment of the households. It was recorded that most of the victims were children and elders who have less role in asset creation and protection.

In the third model, food production was found to be significantly affected by soil fertility problem ( $\beta = -0,314$ ) at less than 1 % probability level. Soil fertility is one of the essential requirements of crop production. Due to overploughing, the land has lost its fertility. Losing this quality of soil creates a huge reduction on the production. Some of the households produced less than the mean expected production as a result of this.

### **INCOME, ASSET, FOOD PRODUCTION AND FOOD AVAILABILITY**

Binary logistic regression was used to show the effect of intermediate variables (food production, income and asset) on food availability (Table 6). The model  $\chi^2$  value was 40,232 and it was significant at less than 1 percent probability level. It also had a prediction success of 75,7 %.

The result showed that asset ( $\text{Exp}(B) = 1,000$ ,  $P < 0,1$ ) and food production ( $\text{Exp}(B) = 1,379$ ,  $P < 0,01$ ) significantly affected food availability of the households. With constant condition of other things, the odd ratio in favoru of food availability increased by 1,000 and 1,379 when the asset possession increased by one Birr and food production increased by one month, respectively.

**Table 5.** Multiple linear regression estimates ( $N = 185$ ).

Independent variables	Model 1: Income		Model 2: Asset		Model 3: Food production	
	Standardized coefficient	t-value	Standardized coefficient	t-value	Standardized coefficient	t-value
(Constant)		5,951		6,403		14,660
Weather Variability	0,006	0,085	-0,016	-0,232	0,042	0,595
Weed	-0,142	-2,116*	-0,081	-1,178	-0,060	-0,844
Plant insect and pest infestation	-0,279	-4,095**	-0,229	-3,283**	-0,075	-1,043
Animal disease and epidemics	-0,065	-0,977	-0,081	-1,184	0,048	0,669
Human disease and epidemics	0,022	0,325	0,145	2,064*	0,114	1,570
Price fluctuation problem	-0,090	-1,305	0,004	0,063	0,011	0,152
Conflict	-0,062	-0,924	-0,054	-0,791	-0,090	-1,264
Soil fertility problem	-0,321	-4,651**	-0,317	-4,489**	-0,314	-4,284**
<b>R<sup>2</sup></b>	0,229		0,190		0,127	
<b>F-test</b>	6,516**		5,169**		3,193**	

\* significant at probability level 1 %

\*\* significant at probability level 5 %

**Table 6.** Logistic regression estimates of factors affecting food availability ( $N = 185$ ).

Variables	B	S.E.	Wald	Sig.	Exp(B)
Assets	0,000	0,000	2,858*	0,091	1,000
Income	0,000	0,000	0,193	0,661	1,000
Food production	0,322	0,076	17,943**	0,000	1,379
Constant	-2,316	0,624	13,776	0,000	0,099
Log likelihood	-89.769				
$\chi^2$	40,232**				
Prediction success	75,7 %				
<b>R<sup>2</sup></b>	0,195 (Cox and Snell), 0,281 (Nagelkerke)				

\* significant at probability level 1 %

\*\* significant at probability level 10 %

The rural households spent most of their food production to cover the need of the family. Thus it is not surprising to see a strong effect of food production on availability. In addition, assets had positive significant effect on food availability. The reason for this is the ability to found food from shared out. Households shared their land and/or livestock and they got some percent of the crop production for their consumption.

## CONCLUSIONS

The study was employed to identify shocks and their effect based on the conceptual framework developed in Ethiopia, Amhara National Regional State of Libo Kemkem District. Through semi structured interview eight shocks were identified as affecting households in the study area. Among these weather variability, weed, plant insect and pest infestation and soil fertility problem were crop shocks. Animal disease and epidemics was an animal shock and price fluctuation problem was a socio economic shock. Human disease and epidemics and conflict were categorized under human shock.

Weather variability creates a mean loss of 3 821,38 Birr while plant insect and pest infestation, weed and animal disease and epidemics creates a mean loss of 886,06 Birr, 508,04 Birr and 1 418,32 Birr respectively. In addition, human disease and epidemics, price fluctuation problem and conflict affect 68,11 %, 88,11 % and 14,59 % of households, respectively.

The results of the study show that 28,1 % of the households were not able to cover their food need while 71,9 % had the ability to cover their food need all over the year. Among the shocks only soil fertility problem had statistically significant relationship ( $p < 0,01$ ) with food availability.

The result of the multiple regression models for income revealed that food production affected by weed existence ( $\beta = -0,142, p < 0,05$ ), plant insect and pest infestation ( $\beta = -0,279, p < 0,01$ ) and soil fertility problem ( $\beta = -0,321, p < 0,01$ ). Asset was found to be affected by plant insect and pest infestation ( $\beta = -0,229, p < 0,01$ ), human disease and epidemics ( $\beta = 0,145, p < 0,05$ ), and soil fertility problem ( $\beta = -0,317, p < 0,01$ ) while food production was affected by soil fertility problem ( $\beta = -0,314, p < 0,01$ ).

The food availability of the households was significantly affected by the asset ( $\text{Exp}(B) = 1,00, p < 0,1$ ) and food production ( $\text{Exp}(B) = 1,379, p < 0,01$ ).

## REMARKS

<sup>1</sup>Human disease and epidemics was measured by the illness score constructed [11].

<sup>2</sup>Birr is the currency of Ethiopia.

## ACKNOWLEDGEMENTS

My first acknowledgement goes to my family, Addisu Damtie, Melkam Addisu, brothers and sister, for their support and encouragement. Also, I would like to thank Dr. Temsegen Tilahun for his valuable comments. My special thanks go to my lovely friends for their quick and kind response when I needed help. Your technical and moral support makes this article real. Finally, I would like to thank Libo Kemkem District Agriculture and Rural Development Office staff members and sample respondents for their cooperation. Thank you all for your help.

## REFERENCES

- [1] Department for International Development: *Defining disaster resilience: a DFID approach paper*. Briefing paper, ACF International. Enhancing resilience to shocks and stresses, 2013,
- [2] Frankenberger, T. and Nelson, S.: *Summary of the Expert Consultation on Resilience Measurement for Food Security*. TANGO International, 2013,
- [3] Department for International Development: *Sustainable Livelihoods Guidance Sheets*. Department for International Development, 1999,

- [4] Ecker, O. and Breisinger, C.: *The Food Security System: A New Conceptual Framework*. IFPRI Discussion Paper 01166, International Food Policy Research Institute, Washington, 2012,
- [5] Headey, D.D. and Kennedy, A.: *Enhancing Resilience in the Horn of Africa, Synthesis of an Evidence Based Workshop*. International Food Policy Research Institute, Washington, 2012, <http://dx.doi.org/10.2499/9780896297977>,
- [6] Disaster Risk Management and Food Security Sector: *Food supply prospects for second half of 2012*. Ministry of Agriculture, Addis Ababa, 2012,
- [7] Berhane, G. et al.: *Evaluation of Ethiopia's Food Security Program: Documenting Progress in the Implementation of the Productive Safety Nets Programme and the Household Asset Building Programme*. International Food Policy Research Institute, 2013,
- [8] Libo Kemkem District Agriculture and Rural Development Office: *Crop production and protection technique dissemination task force plan*. Addis Zemen, 2013,
- [9] Central Statistics Authority: *Summary and statistical report of the 2007 population and housing census. Population size by age and sex*. Addis Ababa, 2008,
- [10] Swaminathan, M.S. and Bhavani, R.V.: *Food production and availability: Essential prerequisites for sustainable food security*. Indian Journal of Medical Research **138**(3), 383-391, 2013,
- [11] Maxwell, D.; Vaitla, B.; Girmay, T. and Rounseville, M.: *Resilience and Livelihoods Change in Tigray, Ethiopia*. Feinstein International Center Research, Tufts University, Somerville, pp.28-30, 2012,

# STUDY OF STAINLESS STEEL RESISTANCE IN CONDITIONS OF TRIBOCORROSION WEAR

Goran Rozing<sup>1, \*</sup>, Vesna Alar<sup>2</sup> and Vlatko Marušić<sup>1</sup>

<sup>1</sup>J. J. Strossmayer University of Osijek, Faculty of Electrical Engineering  
Osijek, Croatia

<sup>2</sup>University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture  
Zagreb, Croatia

DOI: 10.7906/indecs.13.3.10  
Regular article

*Received:* 17 June 2015.  
*Accepted:* 16 July 2015.

## ABSTRACT

Analyzed was the influence of tribocorrosion wear due to effects of fatty acids present in the processed medium. The analysis was conducted on samples made of two austenitic and two martensitic stainless steels. Austenitic steels were tested in their nitrided state and martensitic in their induction hardened state. Conducted were laboratory tests of corrosion resistance of samples, analysis of the microstructure and hardness. To see how the applied processes for modifying the surface of stainless steels behave in realistic conditions, it was conducted the examination of samples/parts of a sunflower cake chain conveyer. Based on the comparison of results obtained in the laboratory and in real conditions, it was estimated that steels AISI 420 and AISI 431 with induction hardened surfaces have a satisfactory resistance to abrasive-adhesive wear in the presence of fatty acids.

## KEY WORDS

tribocorrosion, wear resistance, thermally-chemically treated, austenitic and martensitic steels

## CLASSIFICATION

JEL: L89

PACS: 81.40-z, 81.40.Pq, 81.65.Kn, 81.65.Lp

\*Corresponding author, *η*: goran.rozing@etfos.hr; +385 31 224 689;  
ETFOS, K. Tpimira 2b, HR – 31 000 Osijek, Croatia

## INTRODUCTION

Corrosion adversely affects plant parts and equipment. Direct costs can be seen in terms of the earlier replacement of corroded parts and downtime due to implementation of anti-corrosion protection. Indirect costs are a result of reduction in utilization of installed capacity and/or reduced product quality and thus lower profitability [1]. Failure, or termination of functional performance of parts, usually can appear due to fracture – “breakdown” or as a result of wear, although there are other possible causes. While breakdown is a completely unpredictable – usually instantaneous – process, material wear is mostly gradual, but has a limiting effect on the lifespan of parts. Therefore, one of the ultimate goals of studying the process of wear is to enable parts to function under given conditions with known and acceptable values of friction and wear [2]. Examples of complex wear are especially frequent in the process industry, for example, in food and petrochemical industry. They manifest in combined action of several mechanisms of wear: abrasion, adhesion and corrosion. Therefore, due attention should be given to the study of combined wear. Because, ordinary steels are susceptible to corrosion, and in stainless steels as a problem may occur insufficiently high hardness. In the case of martensitic stainless steels good combination of mechanical properties can be achieved by varying the temperature of austenitising and the duration of annealing [3, 4]. Austenitic stainless steels are widely used because of their excellent corrosion resistance. Wear resistance and surface hardness of austenitic stainless steels is significantly increased in the process of nitriding depending on the duration [5, 6]. Nonetheless, there are limits to the industrial application of this type of steel when exposed to different mechanisms of wear. In this paper we will describe tests that have been carried out in order to compare different corrosion resistant steels exposed to acidic aggressive media in conditions of abrasion-adhesion contact of metal surfaces.

## LABORATORY TESTING

In the experimental part of this study electrochemical corrosion resistance tests were performed on stainless steels, in order to simulate the action of acid aggressive media. These tests are intended for the comparison of corrosion resistance of modified surfaces of stainless steels and were the criteria for selecting materials for conducting tests on samples/parts in real conditions. Tests were conducted on samples of stainless steels: austenitic AISI 304L and AISI 316L and martensitic AISI 420 and AISI 431 (ASTM A240) [7]. Since samples of materials of various physical and chemical properties were used their chemical composition is shown in Table 1.

The contact surfaces of austenitic steels AISI 304L and AISI 316L were modified by a Tenifer process of nitriding. The samples were preheated to 380°C for 3 hours, and then immersed in the salt bath at 580 °C for 5 hours and after that cooled in air. Martensitic steels AISI 420 and AISI 431 were induction hardened with treatment parameters: current frequency  $f = 19$  kHz, austenitization temperature 1050 °C and emulsion quenchant. Wear resistance of these steels is tested on samples/parts in real conditions where they are exposed to abrasion-adhesion wear with the activity of an acidic medium of a pH 5,02 and the presence of abrasive particles of  $\text{SiO}_2 \times n\text{H}_2\text{O}$  of hardness  $\approx 6$  Mohs.

**Table 1.** Chemical composition of tested steels.

Steel	C, %	Mn, %	Si, %	Cr, %	Ni, %	Mo, %
AISI 304L	0,03	1,58	0,41	18,30	8,04	0,07
AISI 316L	0,02	1,49	0,53	17,06	10,22	2,44
AISI 420	0,20	0,46	0,32	12,98	0,20	0,06
AISI 431	0,22	0,42	0,43	15,93	2,03	0,03

## ELECTROCHEMICAL TESTING OF SAMPLES

These tests were conducted in a saturated aqueous solution with CO<sub>2</sub>, the pH value of 5,02 in order to simulate actual conditions of the aggressive environment in which wear of equipment takes place in real conditions. The test samples were made to the dimension of Ø 16×8 mm<sup>2</sup>. Electrochemical tests were conducted in accordance with ASTM G5-94 [8] on the unit Potentiostat/Galvanostat Model 273A EG & E with the use of the software program SoftCorr III. The measurements were made in relation to the reference saturated calomel electrode (SCE) of a known potential of + 0,242 V according to standard hydrogen electrode. Parameters of general corrosion were determined: corrosion potential ( $E_{corr}$ ), corrosion current density ( $j_{corr}$ ), corrosion rate ( $v_{corr}$ ), polarization resistance ( $R_p$ ), pitting potential ( $E_{pit}$ ) and protective pitting potential ( $E_{zpit}$ ).

### Determining the corrosion potential $E_{corr}$

In electrochemical testing the thing that is usually measured is the time change of potential between the electrodes in a galvanic cell through which a weak electric current passes. To determine the electrode potential the electromotive force is measured between the working electrode and a reference electrode of a known and constant potential, which serves as a standard against which to measure the potential of the working electrode. In order to determine the corrosion potential  $E_{corr}$  of tested steels as a function of time, the samples were washed in distilled water and degreased in ethanol. Tests were conducted at room temperature.

### Determining corrosion parameters using the Tafel extrapolation

Electrochemical nature of the corrosion process enables the use of different electrochemical measurement techniques to determine the intensity of corrosion. There are techniques with direct current (DC technique) and techniques with alternating current (AC technique). In electrochemical testing process of corrosion there were used polarization methods of measurement, potentiostatic and galvanostatic polarization. The objective of measuring is to record the polarization current-voltage graph and based on the anode and cathode curves to get a picture of corrosion behaviour of a certain material in a certain medium. Multiple methods for determining the rate of corrosion are based on potentiostatic polarization, such as the Tafel extrapolation and determining the polarization resistance. In this paper potentiodynamic polarization of potential from -250 mV to +250 mV vs  $E_{corr}$ , at a potential change of 5 mV/s, was derived using the Tafel extrapolation method. Results of electrochemical tests of surface-modified samples of selected steels are shown in Table 2.

Results of electrochemical testing shown in Table 2 indicate that the steels AISI 431 and AISI 316L have twice the value of polarization resistance relative to other samples. Values of corrosion current density displayed in Table 2 show that martensitic steels AISI 420 and AISI 431 in their hardened state have the lowest value of corrosion current density among the tested steels. Nitrided austenitic steels have 4 times higher value of corrosion current density.

**Table 2.** Results of electrochemical tests of samples of corrosion resistant steels.

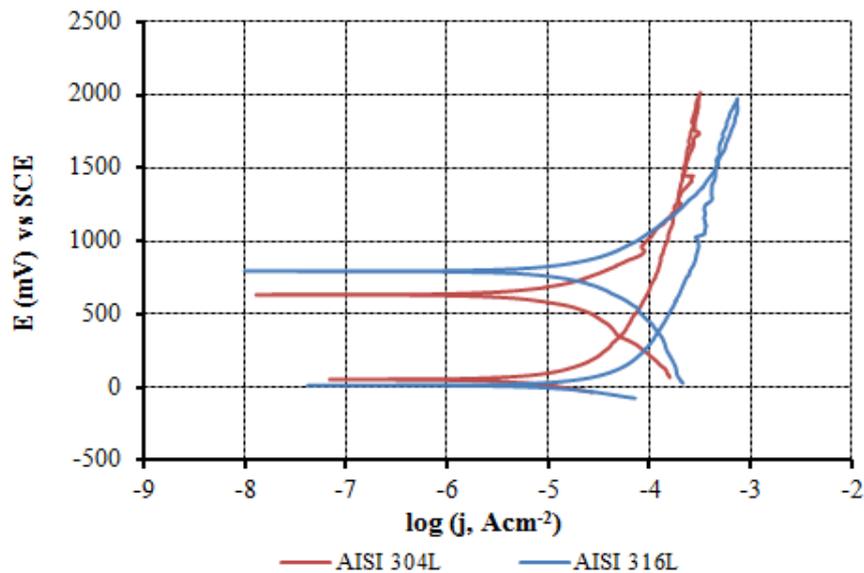
Steel	State	$E_{corr}$ vs SCE, mV	$R_p$ , $\Omega\text{cm}^2$	$\beta_A$ , V/dek	$\beta_K$ , V/dek	$j_{corr}$ , $\mu\text{A}/\text{cm}^2$	$v_{corr}$ , mm/god
AISI 304L	nitrided	-108	3282	0,078	0,103	6,62	0,067
AISI 316L	nitrided	+130	5254	0,988	0,039	6,71	0,069
AISIN420	hardned	-480	3246	0,061	0,083	1,32	0,016
AISI 431	hardned	-212	6766	0,042	0,178	1,65	0,018

The values of corrosion rate for martensitic steels are about the same and 3,3 times less than the ones for samples of nitrated austenitic steels. Generally, less value of corrosion current density and corrosion rate indicates higher corrosion resistance, results in Table 2 indicate that Tenifer procedure in tested case gives less corrosion resistance of austenitic samples in relation to hardened martensitic samples. The cyclic potentiodynamic polarization measurements were carried out in a saturated aqueous solution of  $\text{CO}_2$ , with a pH value of 5,02. Figure 1 shows diagrams of cyclic polarization testing samples of austenitic steels.

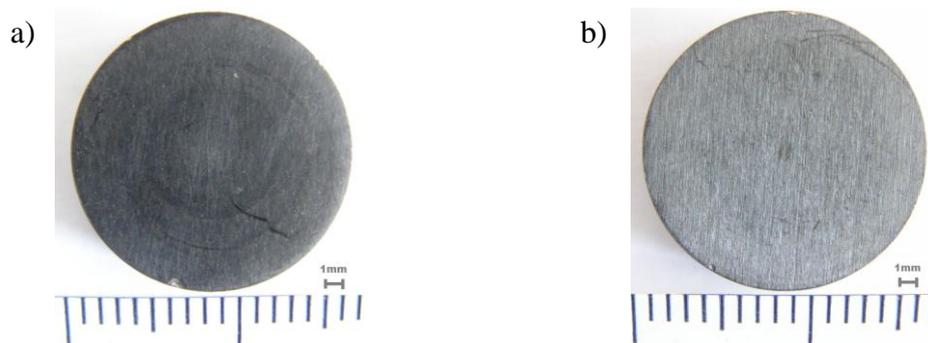
From the diagram of cyclic polarization (Figure 1) it can be seen that samples of nitrated austenitic steels AISI 304L and 316L are not prone to pitting corrosion or crevice corrosion. After recording the curves of cyclic polarization, there were made macro recordings of surfaces of test specimens exposed to electrochemical corrosion. On surfaces of the samples of austenitic steels that were immersed in the electrolyte during the electrochemical tests there are no visible products of corrosion, which is consistent with the literature data [9-11], in Figure 2.

Figure 3 shows overlapping diagrams of cyclic polarization of tested martensitic steels, and Figure 4 macro images of martensitic steel samples after the test.

Diagram of cyclic polarization of martensitic steel samples (Figure 3) shows that the sample of steel AISI 431 shows little inclination top pitting corrosion and crevice corrosion. On the surfaces of samples of martensitic steel that were immersed in the electrolyte during the electrochemical tests there are not visible corrosion products, Figure 4.



**Figure 1.** Diagram of cyclic polarization of austenitic steel samples.



**Figure 2.** The appearance of the surface samples of nitrated austenitic steels after recording a cyclic polarization curve a) AISI 304L, b) AISI 316L.

Tested steels AISI 420 and AISI 431 in their hardened state showed the best corrosion properties both among the hardened martensitic and nitrated austenitic stainless steels. The achieved corrosion resistance represents good prerequisites for the implementation in a real tribosystem, where in addition to mechanical wear the aggressive action of fatty acids is also present. Therefore, martensitic steels were selected for the implementation of wear resistance tests in real conditions.

### MICROSTRUCTURAL ANALYSIS OF MODIFIED SURFACE SAMPLES

Metallographic tests along with hardness measurements provide a complete picture of the material after induction hardening and nitriding. Samples of austenitic stainless steels were corroded with Glycergia and martensitic samples were corroded with solution of HNO<sub>3</sub> and HCl in H<sub>2</sub>O. Microstructure of samples of austenitic stainless steels after nitriding is shown in Figure 5.

Metallographic analysis of samples shows that samples of materials AISI 304L and AISI 316L (Figure 5a and 5b) have an austenitic crystalline structure with strong doubles and finely distributed various precipitates. Along the edges of samples there is a clearly visible nitrated layer onto which continues the austenitic structure of the base material.

The characteristic microstructures of induction hardened martensitic steel samples are shown in Figure 6.

Metallographic analysis of specimens of martensitic steels after surface modification by a method of induction hardening, shows undissolved carbides in the martensitic matrix (Figure 6a) and some retained ferrite in the martensitic matrix (Figure 6b).

### HARDNESS TESTING OF MODIFIED SURFACES OF SAMPLES

Tests of micro-hardness of austenitic steel samples were conducted with a unit DURIMET Leitz, using the method Vickers HV 0,025 for nitrated samples and HV1 for surface hardened samples. Results of hardness testing of nitrated austenitic samples are shown on diagrams in Figure 7, and induction hardened martensitic samples in Figure 8.

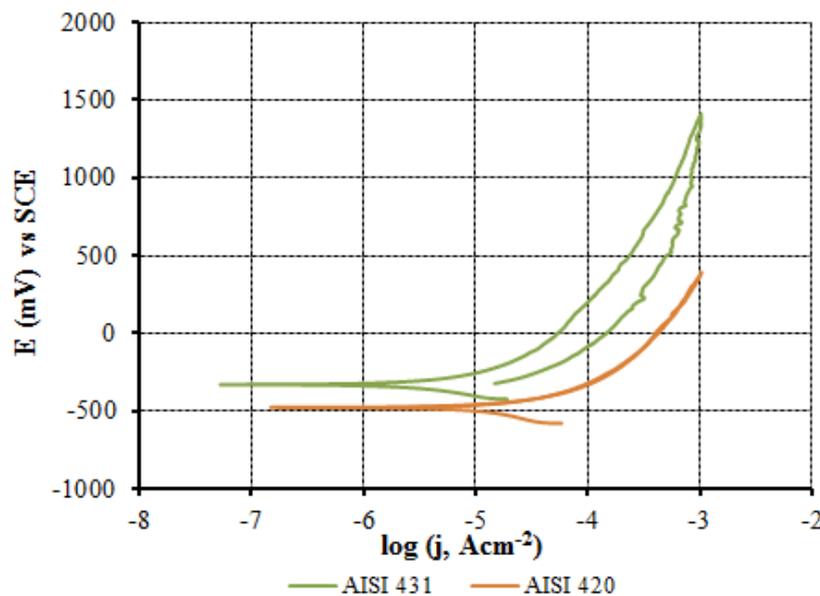
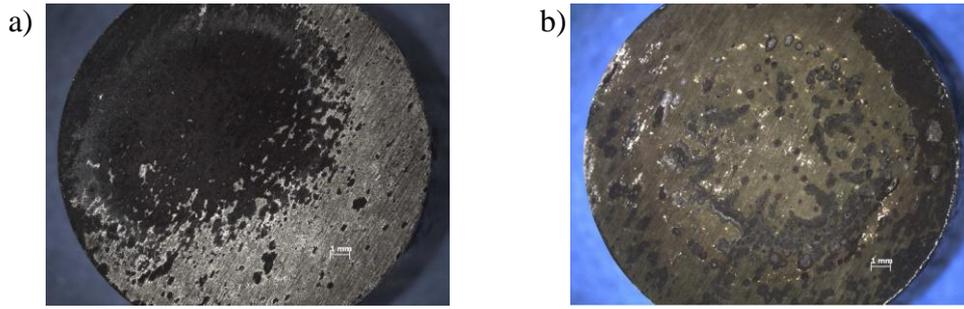
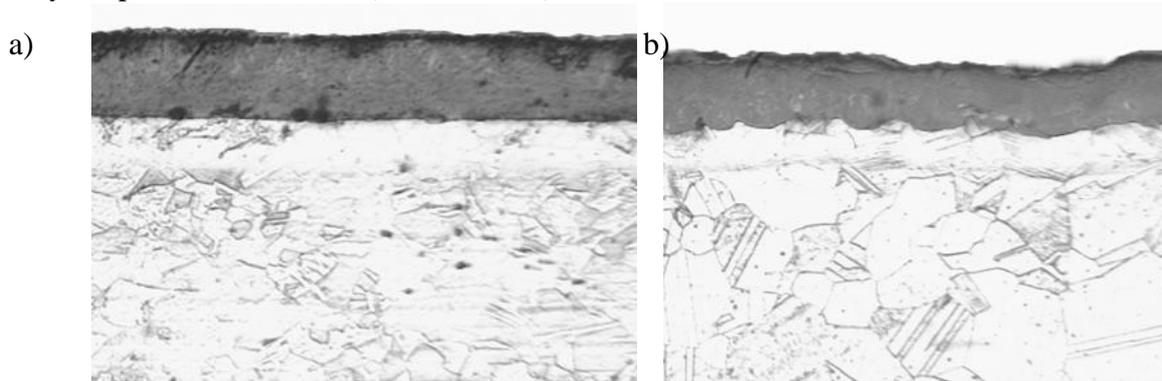


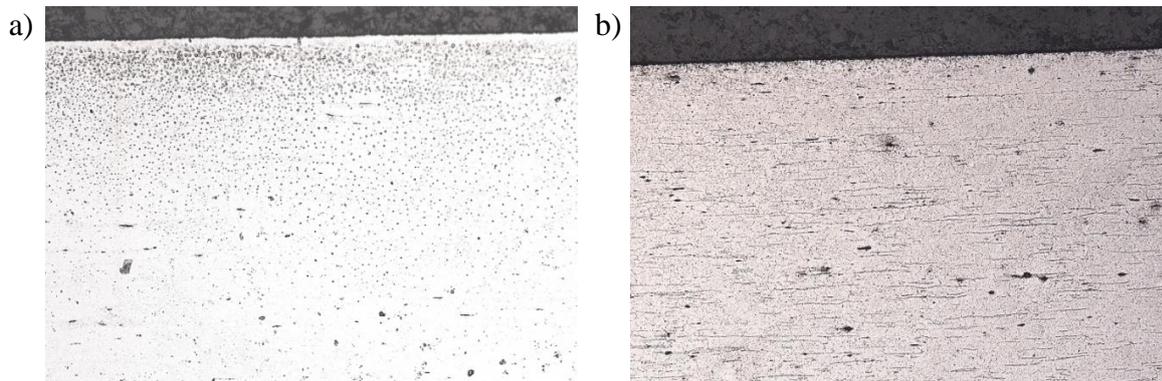
Figure 3. Diagram of cyclic polarization of martensitic steel samples.



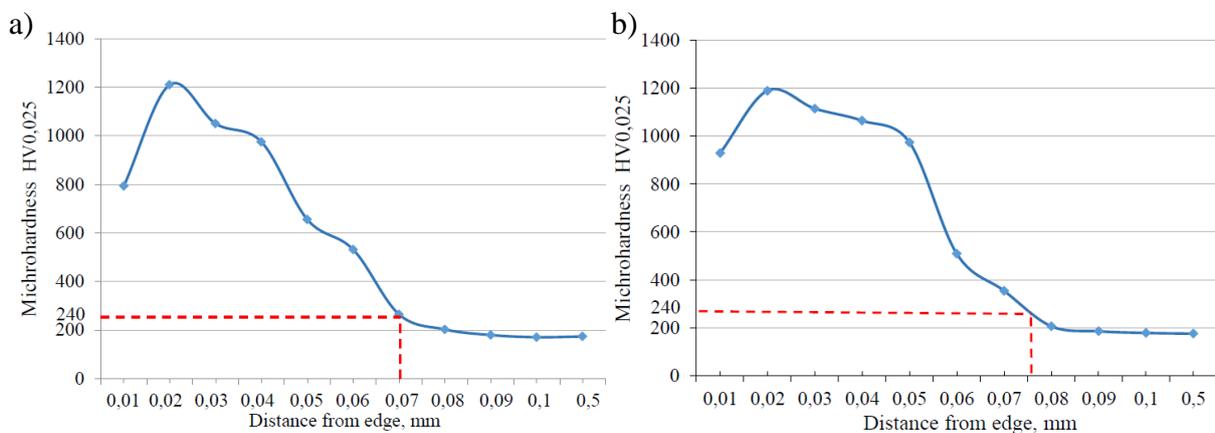
**Figure 4.** The appearance of the surface of hardened martensitic steel samples after recording a cyclic polarization curve a) AISI 420, b) AISI 431.



**Figure 5.** Characteristic microstructure of nitrided steel, magnification of 240:1 a) AISI 304L, b) AISI 316L.



**Figure 6.** Characteristic microstructure of hardened steel, magnification of 200:1 a) AISI 420, b) AISI 431.



**Figure 7.** Hardness curve of nitrided austenitic samples, a) AISI 316L, b) AISI 304L.

The values of the microhardness on cross-sections of samples of AISI 304L and AISI 316L measured from the edge to the core (Figure 7a and 7b) show that the microhardnesses of samples are equal (1200 HV0,025 on the surface) and appropriate for the applied procedure for modifying the surface by nitriding. The effective nitriding depth was  $NHD = 0,07$  mm for nitrided steel AISI 316L, and  $NHD = 0,075$  mm for nitrided steel AISI 304L. Method of Vickers hardness HV1 was used to measure the hardness on cross-sections of induction hardened samples from the edge to the core. The measurement results are shown in the diagrams in Figure 8.

The sample of the steel AISI 420 showed greater surface hardness (610 HV1) relative to the sample of steel AISI 431 which had a surface hardness of (530 HV1), Figure 8. From the hardness curve HV1 according to layer depth measured at the reference line, it is observed the effective depth of surface hardening  $SHD = 2,25$  mm for the sample steel AISI 420 and  $SHD = 1,25$  mm for the steel AISI 431; diagrams in Figure 8.

## TESTING SAMPLES IN REAL CONDITIONS

After the laboratory tests of modified surfaces of stainless steels, martensitic steels AISI 420 and AISI 431 in their induction hardened state were selected for the implementation of wear resistance tests in real conditions. To see the way the surface hardened martensitic steels behave in real conditions of tribocorrosion wear, it was conducted the testing of samples/parts of a sunflower cake chain conveyer, Figure 9.

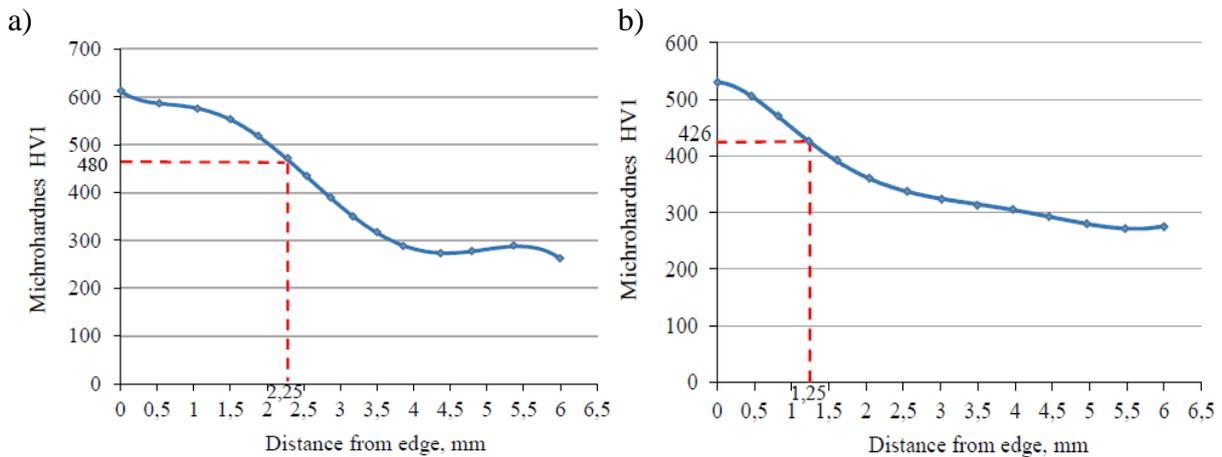


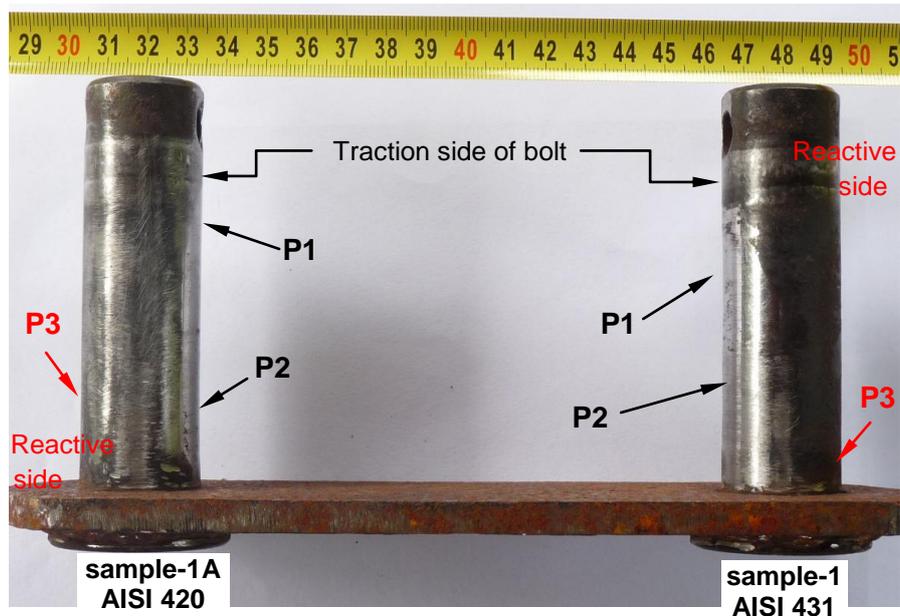
Figure 8. Hardness curve of hardened martensitic samples, a) AISI 420, b) AISI 431.



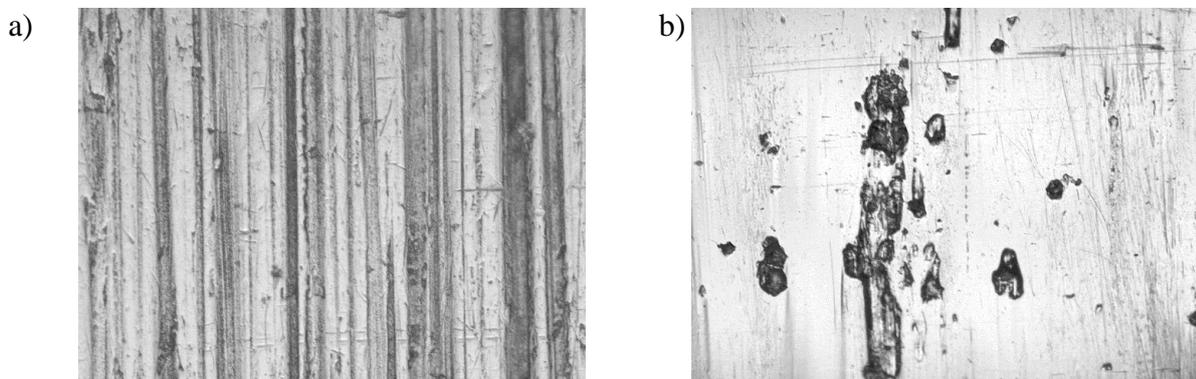
Figure 9. The appearance of the chain of the elevator

During transport of the sunflower cake, the chain of the elevator operates in conditions of fatty acids present in the sunflower cake, which “aggressively” act on the chain parts. In addition to the aggressive action of fatty acids, another cause of wear for the parts of the chain is also the effect of microabrasive particles of  $\text{SiO}_2$  contained in the sunflower cake [9]. Samples of stud bolts were made from selected martensitic steels and built into the chain of the conveyor, and were effectively working for eight months. According to a previous agreement, elevator chain was dismantled and disassembled in order to analyze the wear resistance of test samples of stud bolts. Figure 10 shows a macro image of wear of the sample bolts 1A and 1 in induction hardened state.

In a dimensional control using calipers it was found that there was no construction significant reduction in diameter of bolts. Measured reduction of diameter on induction hardened stud bolts was 0,15 mm for sample 1A (AISI 420) and 0,22 mm for sample 1 (AISI 431). After dimensional control, stud bolts were prepared for recording traces of wear. Recording was performed on a light microscope Leica DM 2500. For the sample of a stud bolt (AISI 420) Figure 11a is showing the traces of wear typical for pure abrasion that is characterized by grooves of different width and depth in the direction of the relative motion of the abrasive/body. Figure 11b is showing the area in which the adhesive wear mechanism is present and where we can see smaller cavities. It is the traction side of the stud bolt 1A, positions P1 and P2.



**Figure 10.** Macro image of wear of stud bolts after tests in use.



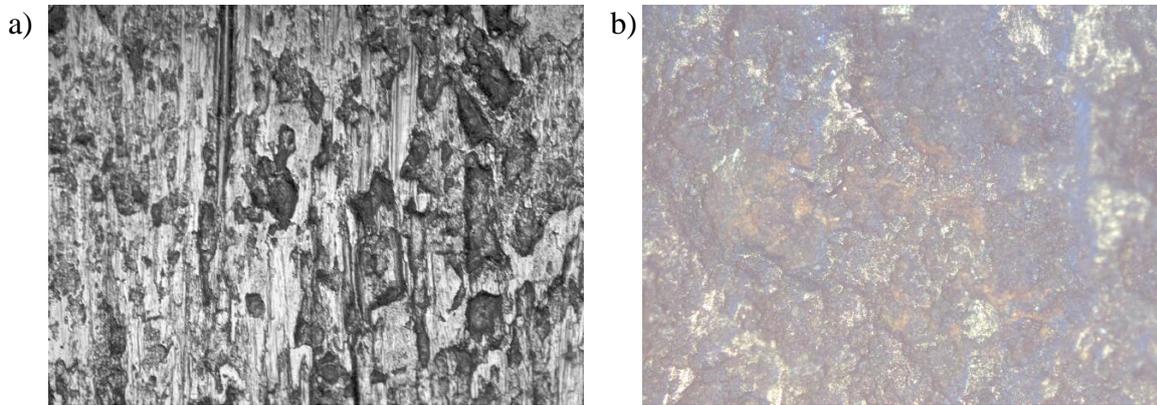
**Figure 11.** Traces of wear of hardened bolts 1A for the steel AISI 420 a) abrasion, position P1, magnification 100:1, b) adhesion, position P2, magnification 50:1.

Figure 12a shows the area with the predominant adhesion mechanism of wear where we can observe cavities of different size and depth, i.e. places where the particles were torn from the surface. It is the traction side of the stud bolt 1 (AISI 431), positions P1 and P2. The part of the stud bolt sample that is not in contact with the sleeve, reveals corrosion products, position P3, Figure 12b.

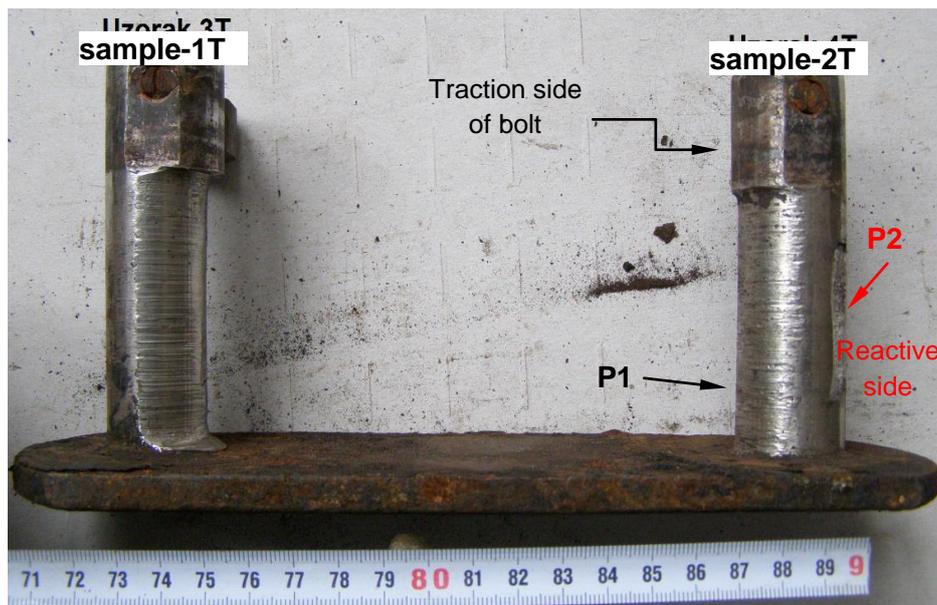
The same mechanisms of wear were determined on surfaces of induction hardened stud bolts made of steels AISI 420 and AISI 431, with no evidence of significant differences in intensity of both abrasion and adhesion wear. The reasons for the appearance of abrasion wear should be sought within the sunflower cake content, more precisely in  $\text{SiO}_2 \times n\text{H}_2\text{O}$  from the sunflower shell as the main carrier of tribological properties. On both parts of the test bolts that are not in contact with the sleeve (position P3) we can notice corrosion products, resulting from the oxidation and/or chemical reactions with aggressive media.

Figure 13 is showing a macro image of worn samples of the stud bolts 1T (AISI 420) and 2T (AISI 431) in a quenched and tempered (as-received) condition.

From Figure 13 it is evident that the stud bolts in a quenched condition have high wear on the traction side, and the dimensional control measured a reduction in diameter of 5 mm.



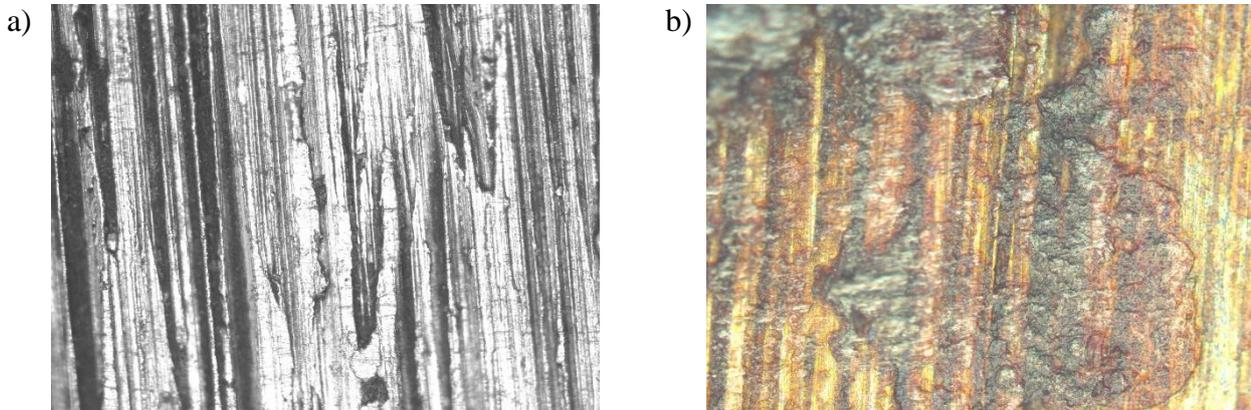
**Figure 12.** Traces of wear of the hardened stud bolt for steel AISI 431, magnification 50:1, a) adhesion + abrasion, position P1, b) corrosion, position P3.



**Figure 13.** Macro image of wear of stud bolts in a quenched condition, 1T – steel AISI 420 and 2T – steel AISI 431.

Figure 14a shows the traces of abrasion wear at the position P1, and in Figure 14b the resulting corrosion products are on the reactive side of the stud 2T, position P2.

Unlike heat-treated stud bolts, stud bolts in a quenched and tempered condition (as-received) have extreme wear on the reactive side.



**Figure 14.** Traces of wear on the traction and reactive side of the stud bolt 4T for AISI 431, a) adhesion + abrasion, position P2, magnification 100: 1, b) corrosion, position P1, magnification 200: 1.

## CONCLUSIONS

Application of corrosion protection in industry today lags behind the theoretical knowledge, which is mainly due to reasons of rapid advances in technology of corrosion protection. Numerous studies have shown that one dollar invested in corrosion protection gives from three to six dollars in terms of longer shelf life and higher reliability of equipment [1].

Given the working conditions of process equipment in the oil industry whose parts are exposed to tribocorrosion activity, their lifespan is shortened and there is system failure. Material losses and maintenance costs are increased. There is a risk of fracture of plant elements, resulting in delays in production. When replacing parts care must be taken about the benefits of materials for producing the parts that are to be installed in equipment in relation to the medium in which the equipment is used. It is necessary to know the mechanical and tribological durability in the selection phase of materials during designs. Strength of materials is based on the use of appropriate coatings during long-term exploitation and can extend the life of equipment parts.

On the basis of tests and analysis of the results obtained, it was concluded that between samples, somewhat greater wear resistance was achieved by steel AISI 420. The achieved greater wear resistance can certainly be attributed to the higher surface hardness and effective hardening depth of 2,25 mm for the surface-hardened steel AISI 420 compared to 1,25 mm for the hardened steel AISI 431. Steel AISI 431 has a slightly higher corrosion resistance due to the higher content of Cr and Ni, but from an economic point of view, taking the price of the base material into account, AISI 420 is up to 50 % cheaper than AISI 431. For this reason it is proposed to use the steel AISI 420 for parts that are used in tribocorrosion conditions that are present in transport of sunflower cake.

## REFERENCES

- [1] Roberge, P.R.: *Handbook of corrosion engineering*. McGraw Hill, 1999,

- [2] Marušić, V.; Rozing, G. and Novoselac, V.: *Actuality of the application of tribological regulators at reducing losses in oil pressing factory.*  
In: Dašić, P., eds.: Proceedings of 10<sup>th</sup> International Conference *Research and Development in Mechanical Industry*. SaTCIP, Donji Milanovac, pp.99-105, 2010,
- [3] Nasery Isfahanya, A.; Saghafiana, H. and Borhanib, G.: *The effect of heat treatment on mechanical properties and corrosion behavior of AISI420 martensitic stainless steell.*  
*Journal of Alloys and Compounds* **509**(9), 3931-3936, 2011,  
<http://dx.doi.org/10.1016/j.jallcom.2010.12.174>,
- [4] López, D. et al.: *Effect of particle velocity and impact angle on the corrosion–erosion of AISI 304 and 420 stainless steels.*  
*Wear* **240**, 80-94, 2005,
- [5] Kumar, T.; Jambulingam, P.; Gopal, M. and Rajadurai, A.: *Surface hardening of AISI 304, 316, 304L and 316L using cyanide free salt bath nitriding process.*  
*International Symposium of Research Students on Materials Science and Engineering*, Chennai, pp.1-11, 2004,
- [6] López, D.; Falleiros, N.A. and Tschiptschin, A.P.: *Effect of nitrogen on the corrosion-erosion synergism in an austenitic stainless steel.*  
*Tribology International* **44**(5), 610-616, 2011,  
<http://dx.doi.org/10.1016/j.triboint.2010.12.013>,
- [7] –: *ASTM A240 Standard Specification for Chromium and Chromium-Nickel Stainless Steel Plate, Sheet, and Strip for Pressure Vessels and for General Applications.*  
ASTM, 2004,
- [8] –: *ASTM G5-94 Standard Reference Test Method for Making Potentiostatic and Potentiodynamic Anodic Polarization Measurements.*  
ASTM, 1999,
- [9] Rozing, G.; Pintarić, A.; Jovanović, V. and Marušić, V.: *Application capability of some steels for chain elements construction exposed to agresively work of media.*  
In: Schauerl, Z. and Šolić, S., eds.: *International conference Matrib 2011*. Croatian Society for Materials and Tribology, Zagreb, pp.411-419, 2011,
- [10] Alar, V.; Juraga, I.; Runje, B. and Alar, Ž.: *Influence of gyoxal on localized corrosion of austenitic stainless steel in spring water.*  
*Materialwissenschaft und Werkstofftechnik* **40**(12), 910-912, 2009,  
<http://dx.doi.org/10.1002/mawe.200900526>,
- [11] Alar, V.; Runje, B. and Alar, Ž.: *Waste Water Treatment after Removal of Thermic Oxides from Stainless Steel Welding Joints.*  
*Srx Engineering* 2010, Id. 261 423, 2010,  
<http://dx.doi.org/10.3814/2010/261423>.

# COMPARISON OF DIFFERENT SIMULATIONS METHODS IN CASE OF SERVICE-PROVIDING COMPANIES

Biserka Runje<sup>1</sup>, Elizabeta Krstić Vukelja<sup>2</sup> and Amalija Horvatić<sup>1</sup>

<sup>1</sup>University of Zagreb – Faculty of Mechanical Engineering and Naval Architecture  
Zagreb, Croatia

<sup>2</sup>University of Zagreb – School of Dental Medicine  
Zagreb, Croatia

DOI: 10.7906/indecs.13.3.11  
Regular article

*Received:* 7 February 2015.  
*Accepted:* 28 April 2015.

## ABSTRACT

Optimal functioning of a market oriented company, in particular the service providing company, is an important example of optimisation of actions within the context of complex environment. In this article we discuss the prospective approach to represent reliably the quality dynamics of such a company, in order to contribute to possible future its quality management.

The agent-based modelling is extracted, among the set of modelling methods, to serve as a frame for representing the generic service providing company and to analyse its dynamics with emphasis on extracting the quality dynamics.

## KEY WORDS

agent based modelling, service providing company, simulation modelling

## CLASSIFICATION

JEL: C63, D21, L22

PACS: 89.65.Gh

## **INTRODUCTION**

Simulation modelling, and numerical simulations in particular, currently figure as a rather useful, in some disciplines unavoidable tool for predicting characteristic elements of a dynamic for the simulated system. That applies to service-providers, the companies which offer specific services on the market. Corresponding examples range from health-service providers, to informatics-communication service providers, to business, to leisure etc.

Their dynamics is highly non-trivial since they ought to operate, in general, within the competitive, market with a broad set of conditions to be fulfilled if they want to be attractive for potential users. Modern companies are often distributed, heterogeneous systems. Their function in providing a service is a result of cooperation among spatially separated elements which work based on different principles, with differences present at all levels, including the subtle ones related to satisfying the different end-user groups.

Among the variety of characteristics of such companies the quality is distinguished as an ultimate, scalar quantity unifying the company's internal characteristics as well as its performance on the market, especially related to end-users experiences. In short, the quality depends on functioning of each and every element of the company. A natural question is how to govern the quality dynamics of a company? In other words, how to modify for a given amount, how to improve, or at least preserve the existing quality level of a company in a competitive, global, rather dynamical market? Typical questions related to quality are, e.g. the following: how to measure quality of the service? How to weigh functioning of company's departments in their contribution to the overall quality? What tolerances should be imposed onto the departments' functioning? What are the proper frequency and intensity in checking the departmental work? And, finally, how to perform regularly and reliably the topics implied in previous questions, so that checking the quality interferes minimally with the regular functioning of the company?

Previously, starting point was the fact that reliably collecting and interpreting the service users attitudes, toward the company and toward the services utilised, is a solid fundament for representing and possibly modifying the dynamics of the company, in particular the dynamics of its quality [1].

In this article we use the simulation modelling in order to reliably represent a given company, or a class of companies, end to enable further researchers or practitioners to utilise the model to check probable company's dynamics depending on the used set of its characteristics.

In section two we list the basics of modelling and simulations, and in section three list and compare several well-known and developed simulation methods. In section four we discuss some implications of choosing the simulation method for service-providing company dynamics simulation.

## **GENERAL CHARACTERISTICS OF MODEL OF SERVICE-PROVIDER**

Let us extract the characteristics of a generic service-providing company, which are encountered in analyses of diverse companies of such type, yet which can be modified in a variable amount depending on a particular case.

Underlying hypothesis is that quality of a service-providing company can be reliably determined (and short-term predicted) based on its accompanying, validated simulation model. Goal of the approach is to test the hypothesis which is conducted directly, i.e. by formulating the simulation model of a service-providing company in a relatively uncharacterised environment. Quality of different parts and the complete model is then evaluated based on the simulation results.

There are several specific goals related to the stated, general goal: (i) quantify the influence of internal information flow, i.e. that within the company, onto the end-user satisfaction and other end-user characteristics, (ii) determine the confidence level of evaluated quality level on the basis of statistical properties of surveyed end-users, (iii) analyse short-term and long-term dynamics of the modelled system, (iv) estimate the intensity of nonlinearity of influences that system's characteristics have onto the end-user satisfaction level, (v) rank micro-level characteristics of system's elements regarding their contribution to the overall end-user satisfaction level, (vi) formulate in details the procedures for service-providers, so that the procedures efficiently contribute to bringing about the predicted change of quality level. Regarding that, the intensity, and sequence of actions involved is to be emphasised. Finally; (vii) extract the characteristic combinations of characteristics of elements of the modelled system as a basis for its validation.

We treat the service-providing company as overall providing the two services and consisting of two departments. Choice of two services is an interpolation between the minimal number of one provided service, and some other number, generally larger than two. With two services one can expect more-complex yet still treatable and interpretable dynamics of service providing. Furthermore, one can expect the ranking of services to emerge as a result of both the users' attitudes and company's own optimisation of formulated goals. One may argue that larger number of services brings about more complicated description but without qualitatively new aspects of dynamics.

Two departments of the company are the department for providing services (DPS) and department for quality management (DQM), the later known also as department for quality control, quality assessment, etc. in various cases. These departments are the part of the same company, yet their functioning is co-opetitive, i.e. simultaneously cooperative and competitive. Being parts of the same firm introduces cooperativeness. However, specific functions put the departments in mutual competitiveness. DPS utilises available resources to perform services to users within the framework as set, at least partially, by the prescriptions from the DQM. On the other hand, DQM measures users satisfaction (to be described in more details further in the text) and formulates prescriptions for DPS in order to achieve optimisation of available resources and achieved user satisfaction. Thus input for DPS are resources from environment and prescriptions from the DQM, while its output are services. Similarly, input for DQM are user attitudes about the services, while its output are prescriptions to DPS. In that way one formulates feedback loop including the two departments and users. The competitiveness between DPS and DQM is seen in handling the prescriptions: one may reasonably expect that DPS solely would like to minimise efforts in providing services from available resources and consequently minimise the influence of prescriptions, especially is they ask for changes in operations of DPS. Conversely, DQM would like the prescriptions to be strict and completely followed, no matter what are the accompanied changes. Formally, one does not need to formulate strictly separated two departments. However, natural division of represented functions and their further implementation in the model is simpler, and easier to interpret is such division is made. The final note regarding the departments is that usually DQM is an existing, single department of a company, but DPS usually consists of several departments, possibly sub-contractors that can be spatially separated. We consider that co-opetiveness of the dynamics of the real departments within DPS is of smaller order of importance than co-opetiveness between DPS and DQM. Because of that we do not refer explicitly to complex structure of the DQM and instead treat it further as a single, monolithic unit.

Along with stated goals and assumed structure of the service-providing company, one cannot formulate the model without simultaneously representing the system's environment in the

model. The environment consists of (i) population (with the service users being a part of it) and of (ii) other institutions.

Population, as part of the modelled environment is represented rather rudimentary, in the form of agents with characteristic dynamics. That dynamics includes internal part, and externally induced change of some of characteristics. One may argue that population's characteristics include age, level denoting the need for the service, as well as other characteristics which are probably of lesser importance for the service itself, but of greater importance for feedback about the provided service. It is assumed that in modern market regularly the feedbacks are collected and utilised in optimisation of services. Regarding feedback collecting, population characteristics include expressivity and satisfaction. For the presented modelling, it suffices to describe the expressivity as a combination of willingness to express ones attitudes, and of intensity of ones attitudes. Satisfaction is equivalent to individual service-user given grade of the service. Collecting and analysing feedbacks is a measurement of quality of services as expressed in part by the users' satisfaction, but screened with their expressivity which introduces a certain bias in the measurement. Since reliability of measurements influences service-providers' dynamics it is to be treated on equal footing as the very providing of services. Preliminary one can think about the following three different populations: younger population with assumed more intense reactions, older population with assumed calmer reactions, and general population serving as a nominal group.

The institutions to be represented in the model are furthermore divided into other companies and into rules which incorporate laws and customs of conducting services. Other companies are either competition, in the assumed competitive market, or collaborators such as contractors or subcontractors. All stated parties have their internal dynamics, as well as mutual interactions.

## **COMPARISON OF SIMULATION METHODS FOR MODELLING SERVICE-PROVIDERS**

Possible simulation methods for stated modelling include system dynamics, neural networks and agent-based modelling. Stated methods are of proven, considerable importance in modelling. However, in the context of service-providing modelling some of their characteristics figure as advantages while other as disadvantages. In the next three sub-sections we describe the three listed modelling methods and their characteristics which are disadvantageous for modelling the service providing company.

Before proceeding, let us state that because of compactness we will not least characteristics that all three simulation methods share, or that are of lesser importance for final decision of their suitability for the present purpose.

### **SYSTEM DYNAMICS**

System dynamics is a powerful method for modelling of diverse types of complex system. It unifies analytic and synthetic approach through representing the modelled system as a combination of elements with relations. The elements are not the non-separable units of matter, energy or information but functionally separated parts of the system, possibly being systems themselves on the next level of modelling.

System dynamics treat modelled system in small number of categories, which generally represent a whole class of entities. In that sense, within a system dynamics model, one would expect to encounter one category representing population, or possibly two or some other small number of categories partitioning the population. However, that may become insufficiently

sensitive to characteristics of an individual user. Since nowadays one tends to use method with the maximal sensitivity to individual user preferences the collective character of categories encountered within the system dynamics is considered as its disadvantage.

## **NEURAL NETWORKS**

Neural network is a method for forming the nonlinear interpolator, the structure of which is based on the initially existing data set relevant to modelled system.

The structure of the neural network, the number of layers, number of neurons in layers and their relations, however, are defined in details based on some statistical properties of interpolated sets as a whole. On the one hand that makes possible formulation of rather simple neural networks. However, on the other hand a drawback occurs in that the resulting structure can be rather independent of realistic structure of modelled system. Since one would like to model a service-provider with fluctuating, or in other way variable, characteristics such as variable number of components and relations, described structure of a neural network figures as its disadvantage for our modelling.

## **AGENT-BASED MODELLING**

Agent-based modelling is simulation method in which functional units are treated as agents, abstract units with prescribed dynamics, e.g. rules of interaction.

That method includes agents, environment and rules. Agents are either individual or collective, regarding the number of humans included in a given agent. In our case, it is opportune that both individual agents (users of services) and collective agents (a company or its departments) are represented in the model.

Main disadvantage of the agent-based modelling is rather large complexity of its validation. However, other stated methods also include non-trivial validation. Therefore we do not treat complex validation of agent-based model as eliminatory property but as an existing disadvantage that must be taken proper care of. Cause of complex validation is related to non-trivial link between the (prescribed) micro-level and obtained macro-level.

## **NOTE ABOUT VALIDATION**

Validation of agent-based models has been separately treated in literature. It is highly non-trivial topic without unique approach. First let us consider approaches to validation of agent-based model in general, and subsequently the validation of agent-based models for services.

## **VALIDATION OF AGENT-BASED MODELS**

Regarding that matter, Liu compares statistical characteristics of data sets generated *in silico* with the data obtained experimentally and in various areas [2]. Validation on such data sets is considered simpler than validation on realistic data, yet its importance is also smaller. Overall, the use of data sets generated in simulations is useful yet not final step in validation of some model [2]. Moss formulated a spectrum of simulation models which can be put in the ordered list so that neighbouring two models differ slightly in relevant characteristics [3]. On the one end of such a list is the empirical model generated based on the realistic data. On the other end of the list is purely theoretical model, possibly modified by internally generated data. Moss assumed that it is possible to formulate a definite relation between the models which are neighbouring in the list, and eventually thereby link the models on the opposite ends of the list [3]. As a result, one could in principle link theory and experiment. Fonoberova et al.

started from the following fact about formulated models: there are certain variations in output quantities caused by variations in input quantities [4]. These variations are in their approach related to pairs of variations that occur in realistic cases. In that sense a dynamical causal relationship is established and utilised as a basis for validation [4].

## **VALIDATION OF AGENT-BASED MODELS FOR SERVICES**

Having in mind the current status of development of validation procedure for agent-based models in general, it is somewhat expected that validation of agent-based models for services is of similar, rather undeveloped status. Contributions to development of that kind of validation include work of Baxter et al. who use agent-based modelling to formulate Customer Relationship Management approach. In that approach the customer population is sufficiently represented [5]. The authors demonstrate the higher quality of their approach in comparison with approaches utilising macroscopic modelling with averaged data about (potential) customers. The modelling they formulate includes customer population dynamics because of what the authors analyse spreading of information about the company among the customer population [5]. Kaihara analyses supply chain in case when the environment (presumably the customer population) is highly dynamic [6]. It is shown that, after taking into account the interactions between the agents on the market (i.e. the customers) one reaches the optimal distribution for allocating the product between the production and selling [6]. Finally, Terano and Naitoh formulate the agent-based model to extract the optimal market strategy in specific, modelled market [7]. The market is characterised with competition among companies. The companies are agents that optimise their behaviour based on the collected, time-dependent data. They validated their results for the markets of television programme and of audio-cassettes [7].

Following described works, we state the hypothesis that it is possible to formulate and validate agent-based model of dynamics of quality in case of innovative company, market service provider. In case the hypothesis is proved then one can use a large number of tools (i.e. the validated models) to simulate reliably and check regularly quality of company's department work, with minimal influence on their predicted functioning. Contrary to that, the rejection of the hypothesis implies either that the contribution of the validated simulation models is rather unimportant to service provider quality dynamics, or that it is not possible to validate formulated models at the present level of understanding. As a result, one would be forced to investigate for possible other approaches to modelling the quality dynamics of a service providing company.

## **CONCLUSION**

To conclude the previous discussion, the problem of quality management is transformed into the following two lower-level problems: (i) how to formulate the simulation model of innovative company, in particular the market service provider, and (ii) how to validate thereby formulated model?

Overall, the stated approach may contribute to development of modelling methodology and simulation of quality management in innovative companies, in particular the service providers.

## **REFERENCES**

- [1] Krstić Vukelja, E. and Runje, B.: *Quality Service Evaluation through the System of Complaints and Praise*. Interdisciplinary Description of Complex Systems **12**(1), 78-91, 2014, <http://dx.doi.org/10.7906/indecs.12.1.5>,

- [2] Liu, F.C.S.: *Validation and Agent-Based Modeling: A Practice of Contrasting Simulation Results with Empirical Data*.  
New Mathematics and Natural Computation **7**(3), 515-542, 2011,  
<http://dx.doi.org/10.1142/S1793005711002050>,
- [3] Moss, S.: *Alternative Approaches to the Empirical Validation of Agent-Based Models*.  
Journal of Artificial Societies and Social Simulation **11**(1), No. 5, 2008,
- [4] Fonoberova, M.; Fonoberov, V.A. and Mezić, I.: *Global sensitivity uncertainty analysis for agent-based models*.  
Reliability Engineering & System Safety **118**, 8-17, 2013,
- [5] Baxter, N.; Collings, D. and Adjali, I.: *Agent-based modelling – intelligent customer relationship management*.  
BT Technology Journal **21**(2), 126-132, 2003,  
<http://dx.doi.org/10.1023/A:1024455405112>,
- [6] Kaihara, T.: *Multi-agent based supply chain modelling with dynamic environment*.  
International Journal of Production Economics **85**(2), 263-269, 2003,  
[http://dx.doi.org/10.1016/S0925-5273\(03\)00114-2](http://dx.doi.org/10.1016/S0925-5273(03)00114-2),
- [7] Terano, T. and Naitoh, K.: *Agent-Based Modeling for Competing Firms*.  
Proceedings of the 37<sup>th</sup> Hawai International Conference on Systems Science, IEEE, 2004,  
<http://dx.doi.org/10.1109/HICSS.2004.1265251>.

# APPLICATION OF INSTRUMENTED CHARPY METHOD IN CHARACTERISATION OF MATERIALS

Željko Alar<sup>1, \*</sup>, Davor Mandić<sup>2</sup>, Andrija Dugorepec<sup>1</sup> and Matija Sakoman<sup>1</sup>

<sup>1</sup>Faculty of Mechanical Engineering and Naval Architecture – University of Zagreb  
Zagreb, Croatia

<sup>2</sup>Research Center for Materials of Region of Istria METRIS  
Pula, Croatia

DOI: 10.7906/indecs.13.3.12  
Regular article

*Received:* 16 May 2015.  
*Accepted:* 29 June 2015.

## ABSTRACT

Testing of absorbed impact energy according to the Charpy method is carried out to determine the behaviour of a material under the impact load. Instrumented Charpy method allows getting the force-displacement curve through the entire test, That curve can be related to force-displacement curve which is obtained by the static tensile test. The purpose of this study was to compare the results of forces obtained by the static tensile test with the forces obtained by the instrumented Charpy method. Experimental part of the work contains testing of the mechanical properties of S275J0 steel by the static tensile test and Impact test on instrumented Charpy pendulum.

## KEY WORDS

impact test, instrumented Charpy pendulum impact test, tensile test

## CLASSIFICATION

JEL: L89, Z00

PACS: 62.20.F-, 81.70.Bt

\*Corresponding author, *η*: zeljko.alar@fsb.hr; +385 1 6168 583;  
Faculty of Mechanical Engineering and Naval Architecture – University of Zagreb, I. Lučića 5,  
HR – 10 000 Zagreb, Croatia

## INTRODUCTION

For most mechanically loaded constructions and parts it is very important to ensure the combination of sufficient strength and toughness. The practice shows that a large number of fractures is not the result of previous plastic deformation, but the consequence of stress which is lower than the yield point. Thus, the concept of toughness is closely linked with the notion of fracture. So we identify brittle and ductile fracture [1].

Material toughness is a property that show the ability to absorb mechanical energy, caused by external, mainly impact energy, through the plastic deformation of materials. The amount of energy consumed for plastic deformation and fracture represents a measure of toughness [2].

Impact test determines the behavior of metallic and polymeric materials in terms of impact loads. Impact test tells us about the energy spented for breaking the specimen whit specific shape and dimensions [1]. Places of stress concentration, such as cracks and notches, are the places where the fatigue, which ends with fracture, begins. It is known that the material fracture in the presence of notch is influenced by the fracture toughness of the material. Therefore, the tests are conducted on specimens with the notch and in this way the multiaxial stress is achived at the root of the notch. Many methods to measure the “strength of the notch” of materials had been developed and standardized. The test is usually carried out at room temperature or at reduced temperatures. Triaxial stress state, the high rate of stress and low temperatures contribute to brittle fracture of materials. Therefore, in order to simulate the heaviest load conditions, almost all tests include the notch which gets broken at various temperatures. The value absorbet impact energy indicates whether the material will break ductile or brittle under the impact stress. Absorbed impact energy is a comparative size and can not be used in calculations when sizing the structural components [3].

## TESTING OF ABSORBED IMPACT ENERGY BY THE INSTRUMENTED CHARPY METHOD

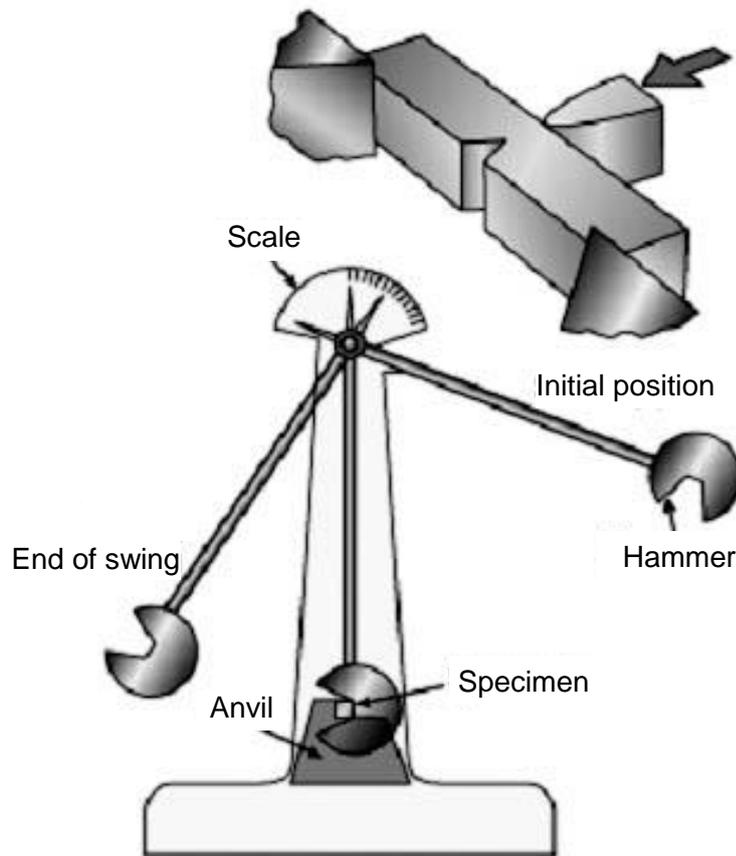
Impact test according the Charpy is carried out to determine the behavior of materials under impact stress. The amount of the absorbed impact energy indicates the “toughness” or “fragility” of material. Charpy impact test is carried out according to the standard EN ISO 148-1 [4]. The energy needed for the fracture of the sample, whose dimensions are defined by the standard, is called absorbed impact energy and it is determined by the following formula [2]:

$$KU(V) = G (h_1 - h_2) \quad (1)$$

where the  $KU(V)$  is absorbed energy (specimens with U or V notch);  $G$  is the weight of hammer;  $h_1$  the initial height of hammer and  $h_2$  the final height of a hammer. Figure 1 shows the schematic view of the test.

It should be noted that the usual Charpy method gives only information with comparative character. The conventional impact test method (without instrumentation) measures the energy required to fracture the specimen under impact loading. For a more complete understanding of the formation and expansion of cracks in the test sample due to impact stress instrumented Charpy method has been developed. Instrumented testing on samples with intentionally made cracks had opened the way for the analysis of fracture mechanics parameters important in the creation and growth of cracks at high loads. The test is carried out according to DIN EN ISO 14556 [6]. The test sample is a Charpy V notch and it is in compliance with EN ISO 148-1.

Instrumented Charpy impact test does not only measures the total energy required to fracture the test specimen, but also the energy at the start of the crack in the notch, energy at maximum

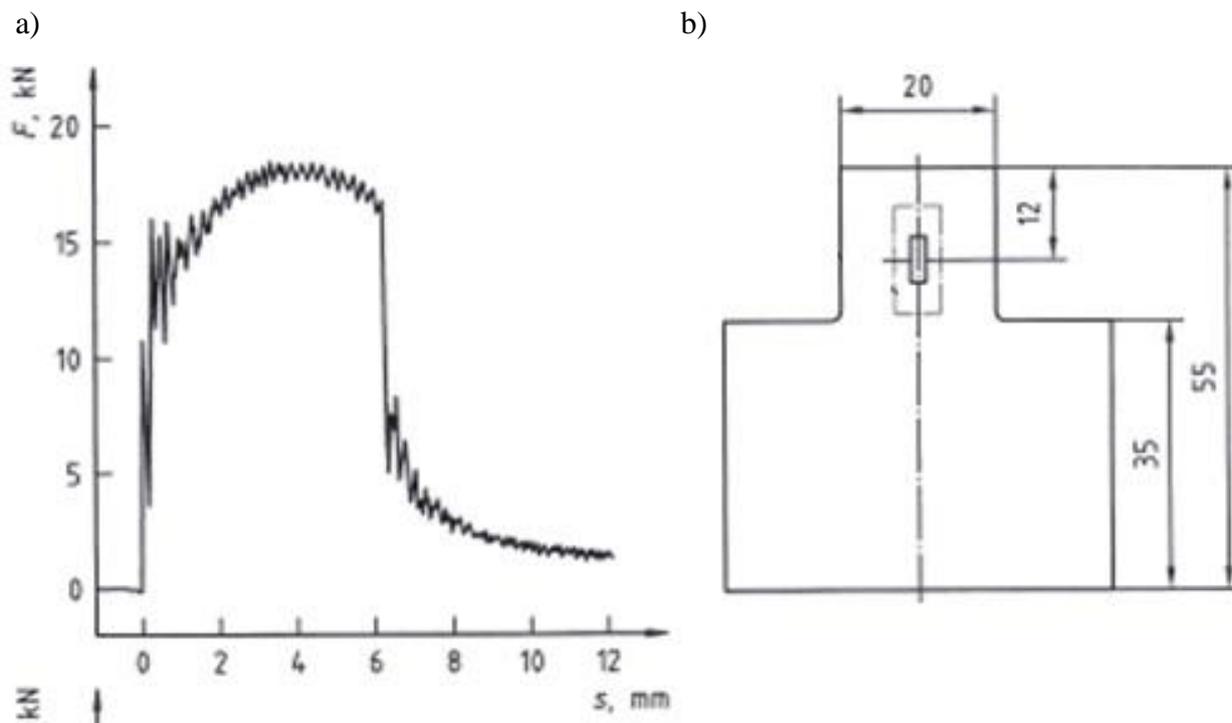


**Figure 1.** Impact testing according the Charpy method [5].

force and energy at the end of unstable fracture. The method requires the use of hammers, which have built-in strain gage to obtain a load curve throughout the entire test procedure. This provides information about the force, displacement, time, and energies that can correlate with mechanical properties such as toughness, fracture toughness KIC, resistance to the fracture, etc. [7].

The advantage of instrumented method is a record in the form of load curves as a function of time by which the absorbed energy during the test can be measured. The method with data on absorbed energy during the test also includes the value of a force versus time. These parameters provide informations about the mechanisms of cracks. Measurement of the force is usually achieved by the use of two active electric resistance strain gauge tensometers at the standard fixed blade to form a force converter. Design of the blade and associated curves are given in Figure 2.

The system for force measurement (instrumented knife, an amplifier, a recording system) must have a response of at least 100 kHz what corresponds to the rise time, no higher than 3.5 ms. Dynamic estimate of the measurement of force can be simplified by measuring the initial value of the first peak. Experimentally, the dynamics of the measuring chain can be considered satisfactory if the steel test specimen with a V notch shows an initial peak higher than 8 kN when the impact velocity is between 5 m/s and 5,5 m/s. This is true if the centers of the active measuring devices are located between 11 mm and 15 mm from the contact point of the hammer. The instrumentation of the blade must be set up so it can produce an appropriate range of force. Instrumented blade must be designed to minimize the vulnerability to asymmetric loads. Experience tells that with V notch samples, nominal impact force appears between 10 kN and 40 kN for all types of steel [6].



**Figure 2.** a) Force-displacement curve, b) sketch of the instrumented knife [6].

## MECHANICAL TESTING OF STEEL S275J0

Mechanical properties of steel S275J0 were tested in the experimental part of the work. Static tensile testing and Impact test on instrumented Charpy hammer with the load-displacement diagrams were carried out. The main objective of the research was to compare the results of forces obtained by the static tensile tests and instrumented Charpy.

### INSTRUMENTED CHARPY IMPACT TEST

Tests were carried out on five specimens with standard sizes of  $55 \times 10 \times 10 \text{ mm}^3$ , V-notched depth of 2 mm, made at angle of  $45^\circ \pm 2^\circ$ . All measurements were within the limit values defined in the standard ISO 148-1. Instrumented Charpy hammer was used with following data (Figure 3): Manufacturer-Zwick/Roell, type: RKP 450, measuring range of device: from 0 J to 450 J, digital readout on PC with resolution of 0,01 J.

The test results are shown in a diagram (Figure 4) for sample 1, which shows that the absorbed impact energy is 185,35 J.

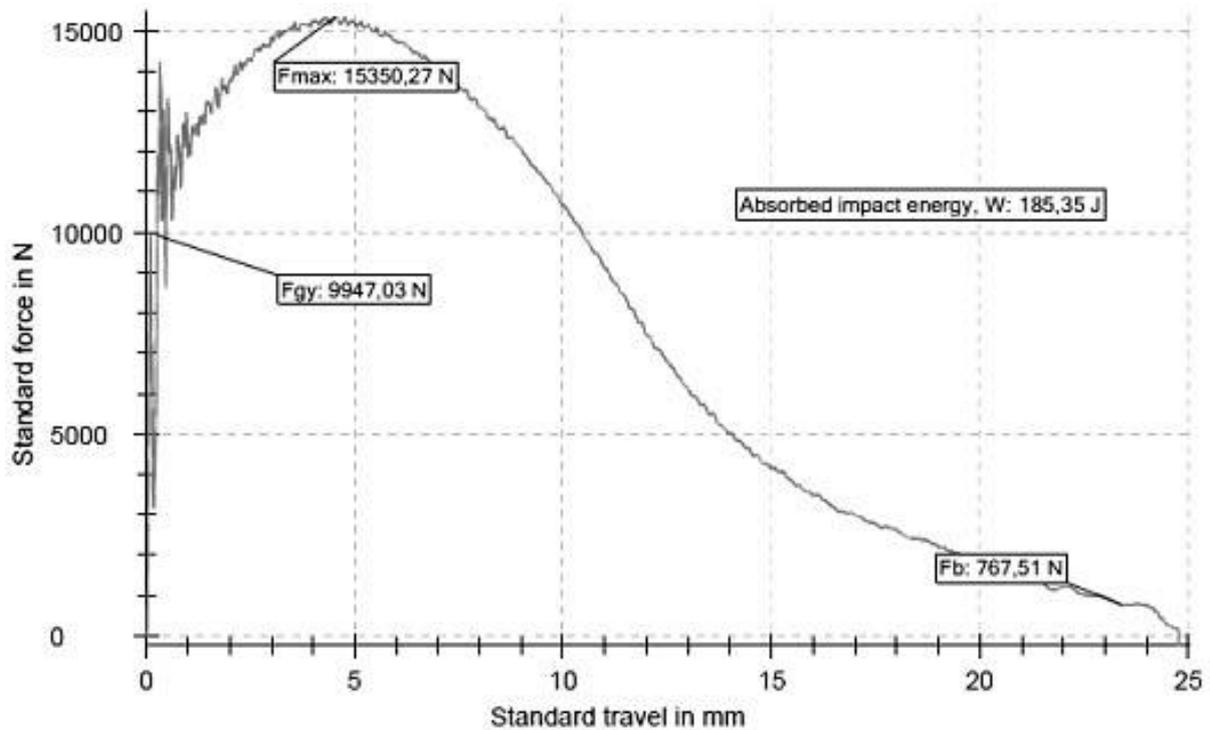
Furthermofre, two different forces are shown in diagram:

- yield force ( $F_{gy}$ ) is determinated on the intersection of the growing part of the second peak of the curve and conlinear regression line through an oscillating curfe force-displacement which amounts 9947,03 N,
- maximum force ( $F_m$ ) is defined as maximum value on the curve, that is 15350,27 N.

Accompanied absorbed impact energies are shown in Table 1.



**Figure 3.** Instrumented Charpy testing machine RKP 450.



**Figure 4.** Results of instrumented Charpy test for sample 1.

**Table 1.** Results of instrumented testing.

Sample	$KV, J$	$F_{gy}, N$	$F_m, N$
1	185,35	9 947,03	15 350,27
2	185,50	8 763,45	15 304,72
3	174,50	9 936,09	15 379,92
4	187,01	10 030,71	15 305,87
5	172,40	9 963,41	15 563,96
<b>Average</b>	<b>180,95</b>	<b>9 728,14</b>	<b>15 380,95</b>
<b>Standard deviation</b>	<b>6,19</b>	<b>483,46</b>	<b>95,79</b>

## STATIC TENSILE TEST

Static tensile test was carried out on a test machine with the following data: Manufacturer: Heckert, type: WPM EU 40 MOD, accuracy class: 0.5 (according to DIN EN ISO 7500-1). The test was carried on five specimens of round cross-section made from the same steel as the specimens for impact test. Specimens are shown in Figure 5.

Results of static tensile tests (maximum force –  $F_m$ , yield force –  $F_{eH}$ , force at break –  $F_k$ , yield strength –  $R_{eH}$ , tensile strength –  $R_m$ , elongation –  $A$ , contraction –  $Z$ ) are shown in Table 2.

The initial diameter of the specimens was 6 mm and the gauge length was 30 mm. Deviations in size and shape of specimens were within the allowable values according to DIN EN ISO 6892-1. The temperature during the test was 22,8 °C, and the loading rate was 5 N/mm<sup>2</sup> s<sup>-1</sup>. Force-elongation diagram of sample 1 is shown in Figure 6.



Figure 5. Specimens for static tensile test.

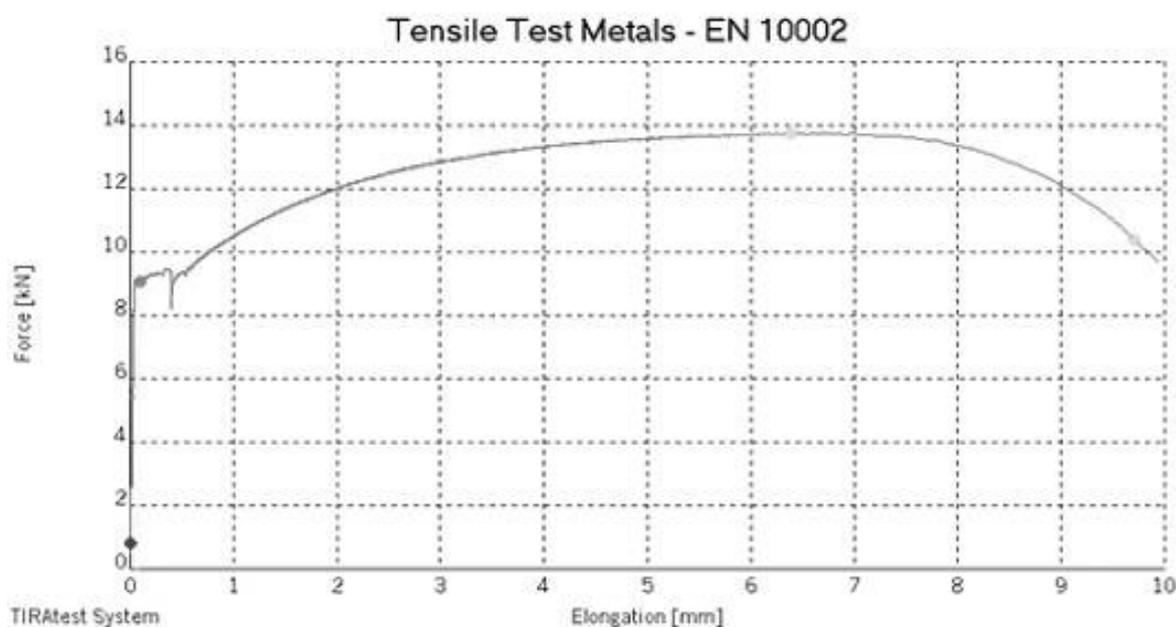


Figure 6. Diagram force-elongation for sample 1.

**Table 2.** Results of static tensile tests.

Sample	$F_{eH}$ kN	$F_m$ kN	$F_k$ kN	$R_{eH}$ N/mm <sup>2</sup>	$R_m$ N/mm <sup>2</sup>	A %	Z %
1	9,37	13,74	10,25	331,40	485,9	34,1	69,2
2	9,57	13,84	10,40	336,15	486,1	37,0	71,6
3	9,50	13,82	10,10	333,90	485,4	36,9	71,7
4	9,47	13,64	10,05	335,17	482,5	35,0	71,6
5	9,32	13,56	9,87	329,80	479,7	36,4	71,6
<b>Average</b>	<b>9,45</b>	<b>13,72</b>	<b>10,13</b>	<b>333,30</b>	<b>483,9</b>	<b>35,9</b>	<b>71,1</b>
<b>Standard deviation</b>	<b>0,09</b>	<b>0,11</b>	<b>0,18</b>	<b>2,36</b>	<b>2,48</b>	<b>1,14</b>	<b>0,97</b>

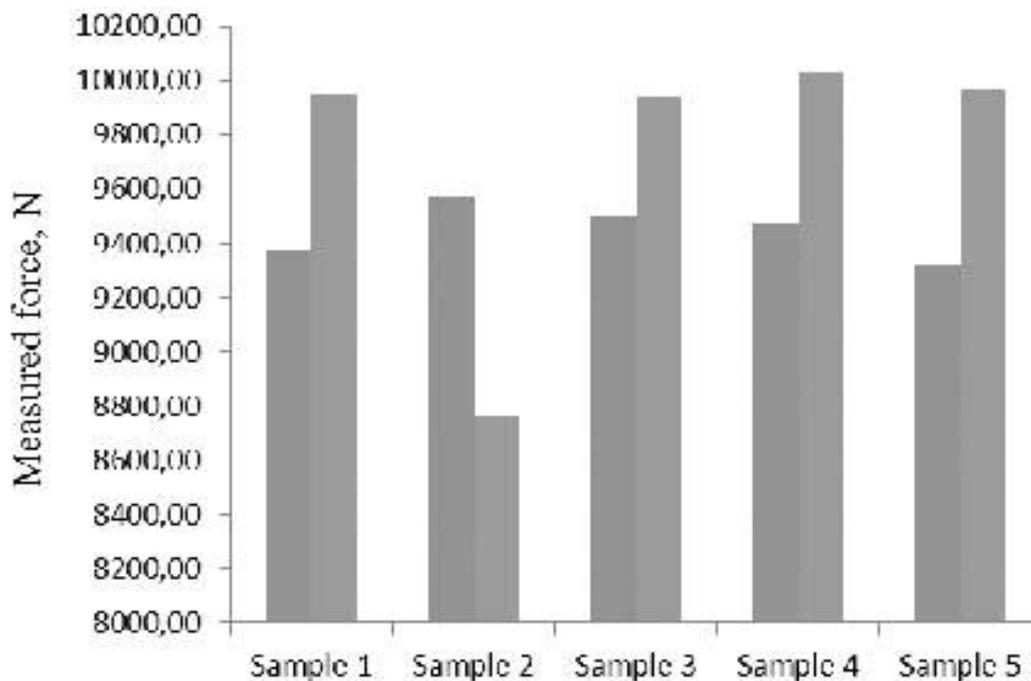
## RESULT ANALYSIS

The purpose of the research was to determine whether there is a correlation between test results obtained by instrumented Charpy method and static tensile testing. By visual comparison of diagrams force-displacement and force-extension (Figures 4 and 6) it can be conclude that there is a correlation between these two diagrams.

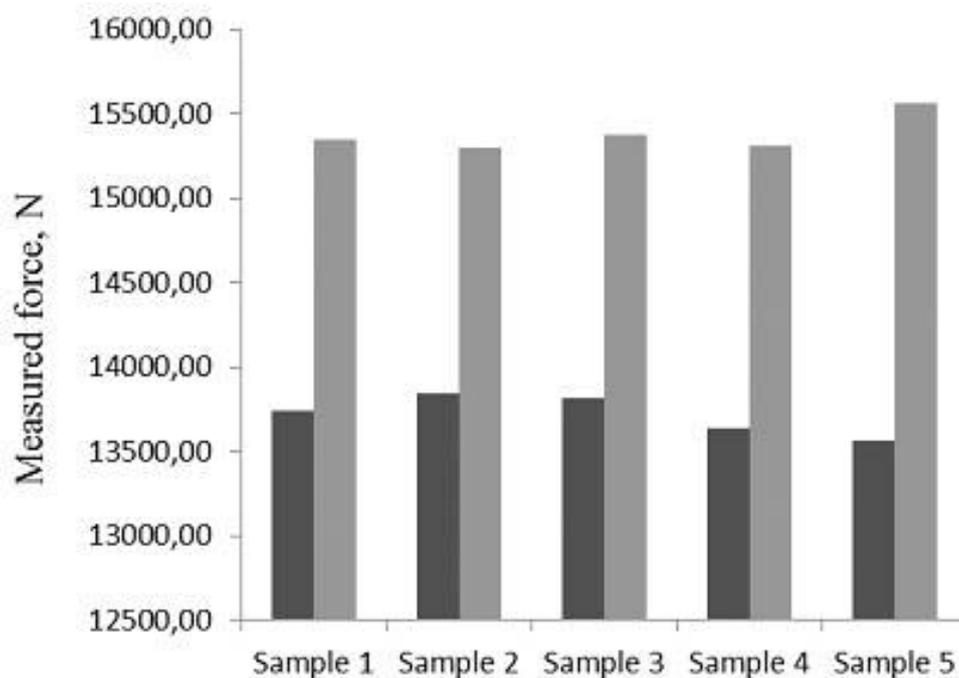
The analysis of the yield force from static tensile test ( $F_{eH}$ ) and yield force from Charpy instrumented test ( $F_{gy}$ ), diagram in Figure 7, shows that these values are approximately equal. A noticeable effect is that the repeatability is better for static tensile test.

The analysis of maximum forces at static tensile test ( $F_m$ ) and the maximum force at Charpy instrumented method ( $F_{m-Ch}$ ) shown in the diagram (Figure 8) It is evident that there is a difference between these values. Reproducibility of results is in this case is much better and it is equal for both test methods.

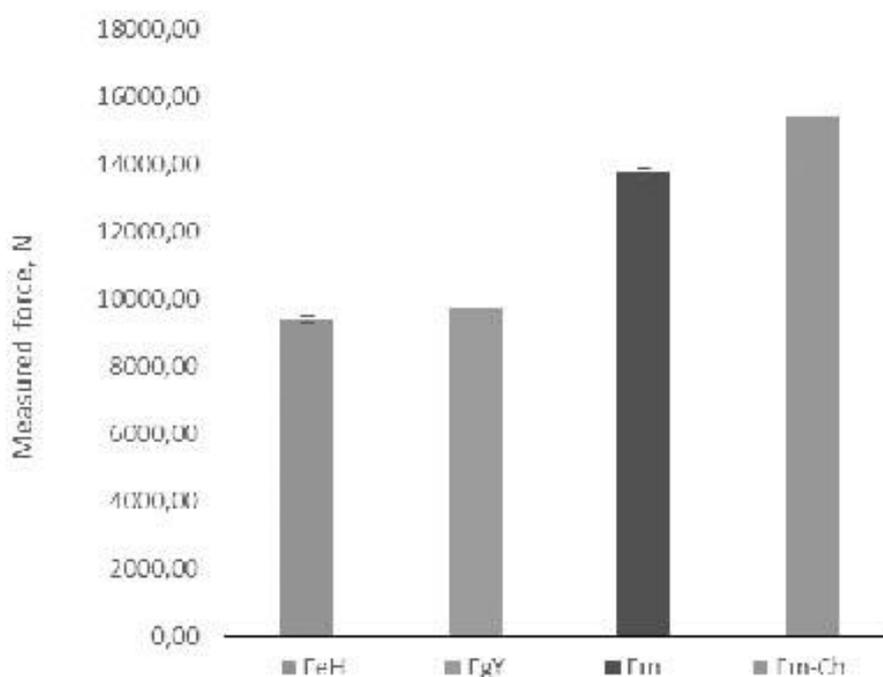
By comparing the arithmetic values of force (Figure 9) it is noticeable that the differences are very small. For the force at yield point, this difference is about 3 % and for the maximum force it is about 12 %.



**Figure 7.** Diagram of the forces  $F_{eH}$  i  $F_{gy}$ . Left bar for a sample denotes  $F_{eH}$  and right  $F_{gy}$ .



**Figure 8.** Diagram of the forces  $F_m$  i  $F_{m-Ch}$ . Left bar for a sample denotes  $F_m$  and right  $F_{m-Ch}$ .



**Figure 9.** Show of the arithmetic values of the force.

Considering that this is a different type of load it is clear that there is some correlation. If all the influential factors were added to testing and evaluation of measurement containing the uncertainty of measurement, the correlation could be confirmed better.

## CONCLUSIONS

With the usage of instrumented Charpy impact test, except the results of absorbed impact energy, more important pieces of information about the material are obtained. Experimentally obtained results make possible reaching certain conclusions.

Based on the force-displacement diagram of the instrumented Charpy method it is plausible to reveal amounts of forces that are characteristic for the static tensile testing. On the basis of tests on S275J0 steel it is shown that yield forces of instrumented method coincide with the values of the forces at the yield point of the static tensile test.

Maximum forces vary around 10 %. Determination of the force at the end of unstable crack and force at the beginning of the crack is problematic because they depend on the type of curves gained from the instrumented Charpy method and in this case it was not possible to connect them with the values from static tensile test. In order to obtain more reliable results, it is necessary to carry out a larger number testing. Additionally samples should be made more uniform, that primarily refers to the angle at V notch with test samples for impact test.

## REFERENCES

- [1] Filetin, T.: *Determining Materials for Product Development*. In Croatian. Faculty of Mechanical Engineering and Naval Architecture, Zagreb, 2000,
- [2] Franz, M.: *Mechanical Characteristics of Materials*. In Croatian. Faculty of Mechanical Engineering and Naval Architecture, Zagreb, 1998,
- [3] Meyers, M. Chawla, K.: *Mechanical Behavior of Materials*. Cambridge University Press, New York, 2009,
- [4] TC 503: HRN EN ISO 148-1:2012 – Metallic materials – Charpy pendulum impact test – Part 1: Test method (ISO 148-1:2009; EN ISO 148-1:2010). In Croatian.
- [5] TWI Ltd.: *Mechanical testing – notched bar or impact testing*. <http://www.twi-global.com/technical-knowledge/job-knowledge/mechanical-testing-notched-bar-or-impact-testing-071>, accessed 15<sup>th</sup> April 2015,
- [6] TC 164: HRN EN ISO 14556: 2008 – Steel – Charpy V-notch pendulum impact test – Instrumented test method, accessed 15<sup>th</sup> April 2015,
- [7] Manahan Sr., M.P.; Cruz Jr., C.A. and John, H.E.: *Instrumented Pendulum Impact Testing of Plastics*. In: Peraro, J.S., ed.: *Limitations of Test Methods for Plastics*. ASTM STP 1369. American Society for Testing and Materials, West Conshohocken, 2000, <http://dx.doi.org/10.1520/stp14346s>.

## TEORIJA DRUŠTVENE PROMJENE KOMPLEKSNOG SUSTAVA: PRIMJENA OPĆEG MODELA 'KRITIČNOSTI'

I. Shimada<sup>1</sup> i T. Koyama<sup>2</sup>

<sup>1</sup>Odsjek za razvojnu biologiju i neuroznanost, Fakultet za znanost o životu, Sveučilište Tohoku

Sendai, Japan

<sup>2</sup>Instituta za istraživanje materijala, Sveučilište Tohoku

Sendai, Japan

### SAŽETAK

Propadanje temelja društva razvijenijih nacija, kako je dramatično demonstrirano kolapsom Lehman, dostiglo je ekstremne razine. Trenutno odlučujuću pazu zauzima stvaranje agenata za promjene. Ovdje konstruiramo model 'kritičnosti' kompleksnog sustava, primjenjujemo ga na opis društvenih promjena te ispitujemo njegovu pouzdanost i validiranost. Model daje eksponencijalnu razdiobu izlaznih društvenih promjena. Validiranost modela provjerena je ispitivanjem udjela glasova političkih stranaka u Japanu. Na temelju rezultata ispitivanja predlažemo novu kvantitativnu strategiju za društvene promjene, "povećavanje informacijske entropije".

### KLJUČNE RIJEČI

društvena promjena kompleksnih sustava, kritičnost, eksponencijalna distribucija, strategija povećanja informacijske entropije, udio glasova

## KONCEPT SLOBODNE VOLJE KAO BESKONAČNE METATEORIJSKE REKURZIJE

H. Hanaan<sup>1</sup> i S. Radhakrishhna<sup>2</sup>

<sup>1</sup>Sibylgaze, Industry Intelligence GmBH  
Bangalore, India

<sup>2</sup>Poornaprajna institut znanstvenih istraživanja  
Sadashivnagar, Bangalore, India

### SAŽETAK

U radu se diskutira kako koncept slobodne volje, slično kao i koncept istine u formalnim jezicima, traži konzistentno definiranu podjelu između razine objekta i meta-razine. Jamesov model dva stanja, koji dekonstruira slobodnu volju u kauzalno otvoreno “slobodno” stanje i njegov okvir u stanju “volje”, neizravan je korak u tom smjeru. Međutim, radi izbjegavanja dileme determinizma, slobodna volja dodatno zahtijeva i beskonačnu regresiju kauzalnih meta-stanja, čime pretvara slobodni izbor u posebni zadatak. Koristimo Jamesov model za definiranje slobodne volje racionalistički usklađene vrste. Time se prirodno postiže razlikovanje kvantne neodređenosti, slobode i slobodne volje, što se redom primjenjuje u umjetnoj inteligenciji, animalnim agentima i ljudskim agentima. U radu zastupamo stav kako kauzalna hijerarhija našeg modela odgovara hijerarhiji Turingove neizračunljivosti. Predlažemo moguće neurobiološke i bihevioralne eksperimentalne provjere slobodne volje. Naznačeno je uokvirivanje modela za fiziku, evolucijsku biologiju, neuroznanost, neuropatološku medicinu i filozofiju morala.

### KLJUČNE RIJEČI

slobodna volja, neizračunljivost, beskonačna rekurzija, Jamesov model dva stanja, kvantna neodređenost

## **SVEMIR JE KAO SFERA. VALNI KONCEPT VREMENA**

A.W. Głuszak i F. Kluza

Sveučilište za znanost o životu  
Lublin, Poljska

### **SAŽETAK**

Postoji prostor za nove ideje o biti vremena. U radu se razmatra naš koncept vremena kao posebna vrsta vala, uz izlaganje rezultata pripadnih istraživanja navedene tematike. Vrijeme definirano kao nosač valova i energije može objasniti mnoge nejasne pojave. Može objasniti gravitaciju, organizaciju našeg planetarnog sustava i brzinu svjetlosti. Iz polazišno pretpostavljenog generiranja mase elementarnih čestica pomoću valova proizlazi hipoteza kako materija postoji zbog gibanja valova vremena. Vrijeme postaje glavna pokretačka sila u svemiru. Razmatrane misli traže daljnju analizu i provjeru a njihovo potvrđivanje može značiti civilizacijsku promjenu.

### **KLJUČNE RIJEČI**

vrijeme, valovi, materija, brzina svjetlosti

## UMJETNOST PAMĆENJA I RAST ZNAJSTVENE METODE

G.P. Sarma

Sveučilište Emory  
Atlanta, Sjedinjene Američke Države

### SAŽETAK

U radu razmatram ideju kako su europske škole sjećanja i pamćenja bile kritične u omogućavanju razvoja znanstvene metode. Nakon povijesnog pregleda razvoja umjetnosti pamćenja, od antičke Grčke do Europe XVII. stoljeća, opisujem kako je Baconovo gledište na znanstvenu metodu bilo fundamentalni dio kulture i šireg dijaloga u kojemu je pamćenje bilo temeljna metodologija strukturiranja znanja i razvoja simboličkih sredstava reprezentiranja znanstvenih koncepta. Značajne ličnosti ovog intenzivnog i brzo razvijajućeg intelektualnog okruženja uključivale su mislioe koje se tradicionalno povezuje sa znanstvenom revolucijom, između ostalih Francisa Bacona, Renea Descartesa i Gottfrieda Leibniza. Rad završavam ispitivanjem akceleriranja matematičke misli u svijetlu umjetnosti pamćenja i njegove uloge u filozofiji XVII. stoljeća, a posebno uloge u Leibnizovom projektu razvoja univerzalnog računanja.

### KLJUČNE RIJEČI

znanstvena metoda, znanstvena revolucija, prosvjetiteljstvo, metodološko promišljanje, univerzalno računanje

## NOVO RAZMATRANJE PISANOG JEZIKA

G.P. Sarma

Sveučilište Emory  
Atlanta, Sjedinjene Američke Države

### SAŽETAK

Velik broj vrhunskih europskih mislioca je, tijekom XVI. i XVII. stoljeća, kako je to povjesničar Paolo Rossi nazvao, bio u potrazi za univerzalnim jezikom, potragom koja je duboko povezana s izviranjem znanstvene metode. IZ moderne perspektive, jedan od mnogih začuđujućih vidova tih napora bio je njihovo oslanjanje na tehnikama pamćenja elemenata. U slučaju Leibnizovog univerzalnog računanja vizija je bila stvoriti slikovni jezik kojeg svatko može naučiti u nekoliko tjedana, a koji bi sadržavao simboličku reprezentaciju svih domena suvremene misli, od prirodoslovlja i teologije do prava. U ovom kraćem radu istražujem zašto je ta potraga bila privlačna misliocima navedenog razdoblja tako što razmatram drevne i moderne podvige pamćenja. Misaonim eksperimentom pokazujem kako društvo sagrađeno u potpunosti na pamćenju može biti manje ograničeno nego nam se inače čini te, nadalje, kako kulturne norme koje potiskuju uporabu pisanog jezika mogu utjecati na razvoj znanstvene metodologije. Gledano na ovaj način, naponi Leibniza i drugih djeluju znatno manje iznenađujuće. Završavam rad općim razmatranjima o interkulturalnim izvorima znanstvene misli.

### KLJUČNE RIJEČI

znanstvena metoda, znanstvena revolucija, pisani jezik, *characteristica universalis*, univerzalno računanje

## **PARADOKS NAGLE I POLAGANE PROMJENE GLEDIŠTA TIJEKOM POLITIČKI REVOLUCIONARNIH VREMENA**

M. Givel

Odsjek za političke znanosti – Sveučilište u Oklahomi  
Norman, Sjedinjene Američke Države

### **SAŽETAK**

Brojni teoretičari političkih revolucija diskutiraju kako su politički revolucionarne aktivnosti dramatične, što rezultira eksplozivnom promjenom orijentacije ustaljenih gledišta zbog čega se javljaju radikalno nova javna gledišta i vladine organizacijske strukture. Ovo istraživanje kvantitativno analizira političke revolucije na vrhuncu u XX. stoljeću, te potvrđuje kako su rijetke kratkoročne revolucionarne aktivnosti koje dovode do uspostavljanja novih gledišta. Najuspješnije političke revolucionarne aktivnosti, koje prate promjene gledišta, su dugoročne. Neke političke aktivnosti nisu bile uspješne. Proces kojim političke aktivnosti zbacuju ustaljena gledišta je kompleksna pojava u kojoj se političke promjene i promjene gledišta odvijaju u znatno razlikujućim vremenskim okvirima.

### **KLJUČNE RIJEČI**

politička revolucija, teorija kompleksnosti, režim gledišta

## NOVIJI ODJECI IZDVOJENIH INDIKATORA RAZVIJENOSTI NA STOPU NEZAPOSLENOSTI: FOKUSIRANJE NA JUŽNOEUROPSKE DRŽAVE

Ksenija Dumcic<sup>1</sup>, Vesna Bucevska<sup>2</sup> i Emina Resic<sup>3</sup>

<sup>1</sup>Sveučilište u Zagrebu – Ekonomski fakultet  
Zagreb, Hrvatska

<sup>2</sup>Sveučilište Sv. Ćirila i Metodija – Ekonomski fakultet  
Skoplje, Makedonija

<sup>3</sup>Sveučilište u Sarajevu – Ekonomski fakultet  
Sarajevo, Bosna i Hercegovina

### SAŽETAK

Rad istražuje relaciju između stope nezaposlenosti i indikatora razvijenosti: (1) BDP po glavi stanovnika u paritetu kupovne moći (izraženom u međunarodnim dolarima) te (2) stopu zastupljenosti interneta definiranu kao postotak korisnika interneta na 100 stanovnika. Za 34 države, u 2013. god. korisnim su se pokazale samo dvije linearne regresije temeljene na prirodnim logaritmima podataka i običnom estimatoru linearne regresije. Jednostavna linearna regresija *Model 1* pokazuje negativnu korelaciju između glavnih proučavanih varijabli  $\ln Y_{UemRate}$  i regresora  $\ln X_{GDPpc}$  te objašnjava otprilike polovicu ukupne varijacije. Jednostavna linearna regresija *Model 2* pokazuje negativnu korelaciju između  $\ln Y_{UemRate}$  i  $\ln X_{IntUse}$ , te objašnjava 27 % ukupne sume kvadrata. Klasteriranje 34 države određeno na temelju tri varijable, primjene Wardove metode i kvadrirane Euklidske udaljenosti daje zanimljivo rješenje s četiri klastera. Fokus je na državama južne Europe, a posebno na državama zapadnog Balkana. Te države nalaze se u tri klastera i nisu homogene. Bosna i Hercegovina te Makedonija u klasteru su sa Španjolskom i Grčkom a sve imaju zahtjevnu ekonomsku situaciju. Albanija, Crna Gora i Srbija u klasteru su s Bugarskom, Rumunjskom i Turskom, kao države jugoistočne Europe. Hrvatska je u klasteru s razvijenijom Italijom, Ciprom i Poljskom te s manje razvijenim Portugalom. Srednje-europska Slovenija pridružena je razvijenijim državama područja, ali najrazvijenije države čine zasebni klaster..

### KLJUČNE RIJEČI

stopa nezaposlenosti, BDP po glavi u paritetu kupovne moći, stopa zastupljenosti interneta, multivarijantna analiza, države zapadnog Balkana

## **RAZLIKOVNI ČIMBENICI I PREPREKE HRVATSKOG IZVOZA U EU15: CAGE PRISTUP**

H. Miloloža

Sveučilište u Zagrebu – Ekonomski fakultet  
Zagreb, Hrvatska

### **SAŽETAK**

Za dugoročno održiv poslovni uspjeh svakog poduzeća, neophodno je promatrati svako tržište zasebno i sustavno se prilagoditi drugačijim uvjetima poslovanja koji vladaju na svakom pojedinom. Međutim, prilagodba može biti kompleksan posao koji iziskuje određene financijske troškove, međutim, ignoriranje razlikovnih kriterija dugoročno dovodi do smanjene profitabilnosti i pada prihoda. U analizi i smanjenju izgubljenog vremenskog odmaka i nepotrebnih troškova uvelike može pomoći primjena CAGE Distance Framework, složenog modela koji obuhvaća sve razlikovne kriterije koji se mogu pojaviti između dvije ili više zemalja.

S obzirom da je recesijom obavijena Hrvatska ušla u Europsku uniju, a još uvijek nije iskoristila dobrobiti koje članstvo pruža niti je u gospodarskom smislu značajno unaprijedila svoje međunarodno poslovanje na jedinstvenom tržištu, očito je da je suočena s poteškoćama u strateškom pristupu pojedinim perspektivnim inozemnim tržištima. Stoga je u ovom radu naglasak stavljen na analizu i identifikaciju razlikovnih faktora između Republike Hrvatske i EU15 zemalja kako bi se olakšala i unaprijedila suradnja među ovim zemljama u budućnosti. Kroz analizu izvoza od nastupa ekonomske krize do danas, predstavljen je i pregled međunarodne razmijene Republike Hrvatske s EU15 te je, primjenom CAGE okvira, napravljena usporedba Republike Hrvatske i svake zemlje EU15 zasebno. Ovom su analizom vidljive značajne oscilacije, neiskorišteni potencijali i problemi hrvatskog gospodarstva u svim dimenzijama CAGE okvira.

### **KLJUČNE RIJEČI**

međunarodno tržište, CAGE okvir, Hrvatska, EU15, izvoz

## UTJECAJ UDARA: EMPIRIJSKA ANALIZA ETIOPIJE

Y.A. Damtie

Sveučilište u Gambelli – Odsjek za krizni menadžment i održivi razvoj  
Gambella, Etiopija

### SAŽETAK

Uz težnju za povećanjem proizvodnje i razvoja također je potrebno smanjiti gubitke uzrokovane udarima. Stanovništvo Etiopije izloženo je prirodnim udarima i udarima ljudskog porijekla. Zbog toga donosioci odluka, vladine i nevladine organizacije trebaju izdvojiti utjecajne udare i njihove učinke koristiti kao polazište svog rada. Istraživanjem su identificirani uzroci nesigurnosti u dobavi hrane i procjenjivanje njihovih učinaka temeljem koncepta razvijenog u okrugu Libo Kemkem regije Amhara u Etiopiji. Podaci su analizirani primjenom deskriptivne statistike, višestruke regresije, binarne logističke regresije te  $\chi^2$ -testai T-testa. Izdvojeno je osam utjecaja kao udara na kućanstva: promjene vremena, količina korova, nametnici na biljkama, problem plodnosti tla, životinjske bolesti i epidemije, ljudske bolesti i epidemije, problem fluktuacije cijena te sukobi. Varijabilnost vremena, nametnici na biljkama, korov, životinjske bolesti i epidemije uzrokovale su prosječni gubitak od 3 821,38 Birr, 886,06 Birr, 508,04 Birr odnosno 1418,32 Birr. Od analiziranih kućanstava 28,1 % nisu mogla zadovoljiti potrebe za hranom tijekom godine dok ih je 71,9 % to moglo. Rezultati višestruke regresije pokazuju značajan utjecaj na dohodak imaju postojanje korova ( $\beta = -0,142$ ,  $p < 0,05$ ), nametnici na biljkama ( $\beta = -0,279$ ,  $p < 0,01$ ) i problem plodnosti tla ( $\beta = -0,321$ ,  $p < 0,01$ ). Značajan utjecaj na imovinu dolazi od nametnika na biljkama ( $\beta = -0,229$ ,  $p < 0,01$ ), ljudskih bolesti i epidemija ( $\beta = 0,145$ ,  $p < 0,05$ ) te problema plodnosti tla ( $\beta = -0,317$ ,  $p < 0,01$ ) dok na proizvodnju hrane značajno utječe problem plodnosti tla ( $\beta = -0,314$ ,  $p < 0,01$ ). Model binarne logističke regresije pokazuje kako je dostupnost hrane u kućanstvima značajno ovisna o imovini ( $\text{Exp}(B) = 1,00$ ,  $p < 0,1$ ) i proizvodnji hrane ( $\text{Exp}(B) = 1,379$ ,  $p < 0,01$ ).

### KLJUČNE RIJEČI

udari, utjecaj, Etiopija, dostupnost hrane, binarno logistička regresija, višestruka regresija

## ISTRAŽIVANJE OTPORNOSTI KOROZIJSKI POSTOJANIH ČELIKA U UVJETIMA TRIBOKOROZIJSKOG TROŠENJA

Goran Rozing<sup>1</sup>, Vesna Alar<sup>2</sup> i Vlatko Marušić<sup>1</sup>

<sup>1</sup>Sveučilište J. J. Strossmayera u Osijeku – Elektrotehnički fakultet Osijek  
Osijek, Hrvatska

<sup>2</sup>Sveučilište u Zagrebu – Fakultet strojarstva i brodogradnje  
Zagreb, Hrvatska

### SAŽETAK

Analiziran je utjecaj tribokorozijskog trošenja usljed djelovanja masnih kiselina prisutnih u prerađivanom mediju. Analiza je provedena na uzorcima izrađenima od dva austenitna i dva martenzitna korozivski postojana čelika. Austenitni čelici su ispitani u nitriranom stanju, a martenzitni u indukcijski zakaljenom. Provedena su laboratorijska ispitivanja korozivske otpornosti uzoraka, analiza mikrostrukture i tvrdoća. Kako bi se vidjelo na koji se način primjenjeni postupci modificiranja površina korozivski postojanih čelika ponašaju u realnim uvjetima, napravljeno je ispitivanje uzoraka/dijelova na lančastom transporteru suncokretove pogače. Na osnovu usporedbe rezultata ispitivanja u laboratoriju i u realnim uvjetima, procijenjeno je da za uvjete abrazijsko-adhezijskog trošenja uz prisustvo masnih kiselina, zadovoljavajuću otpornost imaju čelici AISI 420 i AISI 431 s indukcijski zakaljenim površinama.

### KLJUČNE RIJEČI

tribokorozija, otpornost trošenju, toplinsko-kemijski obrađeni, austenitni i martenzitni čelici

## **USPOREDBA RAZLIČITIH METODA SIMULACIJE NA PRIMJERU TVRTKI KOJE PRUŽAJU USLUGE**

B. Runje<sup>1</sup>, E. Krstić Vukelja<sup>2</sup> i A. Horvatić<sup>1</sup>

<sup>1</sup>Sveučilište u Zagrebu – Fakultet strojarstva i brodogradnje  
Zagreb, Hrvatska

<sup>2</sup>Sveučilište u Zagrebu – Stomatološki fakultet  
Zagreb, Hrvatska

### **SAŽETAK**

Optimalno djelovanje tržišno orijentirane tvrtke, posebno one koja nudi usluge, značajan je primjer optimiranja akcija u kompleksnoj okolini. U ovom radu diskutiramo moguće pristupe za pouzdano predstavljanje dinamike kvalitete takvih kompanija, kako bismo doprinijeli njihovom mogućem, budućem upravljanju kvalitetom.

Među metodama modeliranja, izdvojena je metoda modeliranja pomoću agenata, kako bi poslužila kao okvir za predstavljanje generičke tvrtke koja pruža usluge te za analiziranje njene dinamike s naglaskom na izdvajanju dinamike kvalitete.

### **KLJUČNE RIJEČI**

modeliranje pomoću agenata, tvrtka koja pruža usluge, simulacijsko modeliranje

## **PRIMJENA INSTRUMENTIRANOG CHARPYEVOG BATA ZA KARAKTERIZACIJU MATERIJALA**

Željko Alar<sup>1</sup>, Davor Mandić<sup>2</sup>, Andrija Dugorepec<sup>1</sup> i Matija Sakoman<sup>1</sup>

<sup>1</sup>Sveučilište u Zagrebu – Fakultet strojarstva i brodogradnje  
Zagreb, Hrvatska

<sup>2</sup>METRIS – Centar za istraživanje materijala Istarske županije  
Pula, Hrvatska

### **SAŽETAK**

Provedeno je ispitivanje apsorbirane energije metodom Charpyevog bata zbog karakterizacije materijala pri udarnom radu loma. Metoda instrumentiranog Charpyevog bata omogućava određivanje krivulje sile u ovisnosti o pomaku za cijelo ispitivanje. Ta krivulja može se povezati sa krivuljom sile u ovisnosti o pomaku pri statičnom vlačnom ispitivanju. Cilj ispitivanja je uspoređivanje rezultata za iznose sila dobivenih navedenim dvjema metodama. Eksperimentalni dio rada uključivao je određivanje mehaničkih svojstava čelika S275J0 statičkim vlačnim ispitivanjem i ispitivanjem instrumentiranim Charpyevim batom.

### **KLJUČNE RIJEČI**

udarni rad loma, instrumentirani Charpyev bat, vlačno ispitivanje



## MANUSCRIPT PREPARATION GUIDELINES

Manuscript sent should contain these elements in the following order: title, name(s) and surname(s) of author(s), affiliation(s), summary, key words, classification, manuscript text, references. Sections acknowledgments and remarks are optional. If present, position them right before the references.

**ABSTRACT** Concisely and clearly written, approx. 250 words.

**KEY WORDS** Not more than 5 key words, as accurate and precise as possible.

**CLASSIFICATION** Suggest at least one classification using documented schemes, e.g., ACM, APA, JEL, PACS.

**TEXT** Write using UK spelling of English. Preferred file format is Microsoft Word. Provide manuscripts in grey tone. For online version, manuscripts with coloured textual and graphic material are admissible. Consult editors for details.

Use Arial font for titles: 14pt bold capital letters for titles of sections, 12pt bold capitals for titles of subsections and 12pt bold letters for those of sub-subsections.

Include figures and tables in the preferred position in text. Alternatively, put them in different locations, but state where a particular figure or table should be included. Enumerate them separately using Arabic numerals, strictly following the order they are introduced in the text. Reference figures and tables completely, e.g., “as is shown on Figure 1, y depends on x ...”, or in shortened form using parentheses, e.g., “the y dependence on x shows (Fig. 1) that...”.

Enumerate formulas consecutively using Arabic numerals. In text, refer to a formula by noting its number in parentheses, e.g. formula (1). Use regular font to write names of functions, particular symbols and indices (i.e.  $\sin$  and not *sin*, differential as  $d$  not as *d*, imaginary unit as  $i$  and not as *i*, base of natural logarithms as  $e$  and not as *e*,  $x_n$  and not *x<sub>n</sub>*). Use italics for symbols introduced, e.g.  $f(x)$ . Use brackets and parentheses, e.g.  $\{[( )]\}$ . Use bold letters for vectors and regular GoudyHandtooled BT font (for MS Windows) or similar font for matrices. Put 3pt of space above and below the formulas.

Symbols, abbreviations and other notation that requires explanation should be described in the text, close to the place of first use. Avoid separate lists for that purpose.

Denote footnotes in the text by using Arabic numerals as superscripts. Provide their description in separate section after the concluding section.

References are listed at the end of the article in order of appearance in the text, in formats described below. Data for printed and electronic references is required. Quote references using brackets, e.g. [1], and include multiple references in a single bracket, e.g. [1-3]. Mention all authors if there are not more than four of them, starting with surname, and followed with initial(s), as shown below. In other cases mention only the first author and refer to others using et al. If there are two or more authors, separate the last one with the word “and”; for other separations use semicolon. Indicate the titles of all articles, books and other material in italics. Indicate if language is not English. For other data use 11pt font. If both printed version and the Internet source exist, mention them in separate lines. For printed journal articles include journal title, volume, issue (in parentheses), starting and ending page, and year of publication. For other materials include all data enabling one to locate the source. Use the following forms:

- [1] Surname, Initial1.Initial2.; Surname, Initial1.Initial2. and Surname, Initial1.Initial2.: *Article title*.  
Journal name **Vol**(issue), from-to, year,  
<http://www.address>,
- [2] Surname, Initial1.Initial2. and Surname, Initial1.Initial2.: *Book title*.  
Publisher, city, year,
- [3] Surname, Initial1.Initial2.; Surname, Initial1.Initial2., eds.: *Title*.  
Conference title, dates from-to year. Publisher, city, year of Proceedings publishing.

If possible, utilise the template available from the INDECS web page.

**CORRESPONDENCE** Write the corresponding author’s e-mail address, telephone and address (i.e.,  $\eta$ ).

**ISSN 1334-4684 (printed)**  
**<http://indecs.eu>**