CONTENTS

⬢ Research

▲ Experiences & Tools

A quali-quantitative study on the definition of types of non-hospital residential facilities for psychiatric patients

Massimo Miglioretti[1], Augusto Monge Roffarello[2], Marcella Ercole[2], Fabrizio Zucca[2]

[1] Department of Psychology, University of Milano-Bicocca, Milan, Italy
[2] Studio Co.S.S. – Consulting & Services for Health, Turin, Italy

ABSTRACT. Questo studio si proposto di identificare alcune tipologie di strutture residenziali per pazienti psichiatrici presenti sul territorio e verificare se vi siano differenze tra i pazienti che vivono in ciascuna di queste. Lo studio, quali-quantitativo, ha permesso di identificare 8 tipologie di strutture. Ciascuna di esse ospitava pazienti psichiatrici con caratteristiche significativamente differenti. Tale classificazione delle strutture residenziali per pazienti psichiatrici può essere utile in sede di programmazione sanitaria in quanto mette in luce i diversi approcci di cura che vengono utilizzati negli interventi di comunità e l'identità specifica di ogni approccio.

SUMMARY. The de-institutionalisation of psychiatric patients has led to the construction of various forms of residential facilities for people with mental illness in the community. This study had two aims: to identify the types of residential facilities for psychiatric patients and to determine whether there are differences between patients who live in different types of these. A mixed method approach was used. Interviews with the managers of 13 residential facilities were carried out. Quantitative data about the environmental characteristics, human resources, and characteristics of patients recovered of each residential facility were collected. We identified 8 types of facilities. The characteristics of the psychiatric patients of each residential facility were significantly different from those of the patients of the others. The classification of residential facilities for psychiatric patients is useful for describing different approaches to care that are used in community interventions and the identify the specificity of each approaches.

Keywords: *Residential facilities, Psychiatric patients, Communities treatment, Mixed method approach*

BACKGROUND

Reviews of the systems of psychiatric services that began mainly in the second part of the last century in Italy and, more generally, in Europe have led to the goal of de-institutionalisation. Several studies have shown that the dissolution of the asylum system has favoured the emergence of various residential solutions that are more or less integrated with the community and offer different types of assistance and rehabilitative or therapeutic possibilities (Brunt & Hansson, 2002; de Girolamo et al., 2007; Fakhoury, Murray, Shepherd & Priebe, 2002; Lora, 2009). Most research on this subject has highlighted how these solutions have effectively responded to the different needs of users, although questions remain about the quality of care offered and the integration of these solutions into a coherent system (de Girolamo et al., 2007; Morris, Lora, McBain & Saxena, 2012; Thornicroft & Tansella, 2004).

In recent years in Europe, some attention has been paid to how targeted structures for the assistance, care and rehabilitation of patients with psychiatric disorders have been developing. The DEMoBinc project (Development of a European Measure of Best Practice for People with Long Term Mental Illness in Institutional Care) is a striking example. This project is aimed at building and validating an appropriate tool to assess the living conditions and care quality provided to patients enrolled in psychiatric residential facilities (Taylor et al., 2009). This project has led to the development of the QuIRC (Quality Indicator for Rehabilitative Care), which was the first tool to compare, in the very diverse context of European psychiatric care, aspects of the quality of services provided by residential facilities for psychiatric patients (Killaspy et al., 2011).

In Italy, the PROGRES research programme was intended to carry out the first survey of, and an initial evaluation of, the complete range of assistance facilities that arose after the approval of Law 180, which mandated deinstitutionalisation and disclosure by psychiatric hospitals. This research is unique in its descriptive power and the quality of the data acquired and has highlighted several positive aspects of residential facilities and their operation, but it has also clearly left some gaps; for example, the current research has failed to ascribe any therapeutic-rehabilitative qualities to these facilities, which are described as "homes for life" rather than transitional support (Picardi, de Girolamo & Morosini, 2003). This research has shown that, in Italy, heterogeneous residential facilities for psychiatric patients have been developed. Currently, a classification system is needed to account for both the services that these facilities provide and the needs of the users of these facilities (Santone et al. 2005).

This study was developed in light of existing questions about the definitions of appropriate criteria for the indication of treatment, care and rehabilitation programs for residential facilities for psychiatric patients that have arisen in the Italian context. The present work had two primary, inter-connected aims. The first was to identify the different types of residential facilities for psychiatric patients; this process began with an analysis of the services that each facility provided and continued by defining aspects of the methods, assistance processes, treatments and rehabilitation of the guests of each residential facility type. Second, this research sought to identify differences between patients who live in each type of residential facility identified.

METHOD

Participants

Five private companies took part in this research; some of these companies owned a single residential facility, and others owned and/or operated multiple residential facilities in both the Piedmont or, more generally, in Northern Italy. Seven semi-structured interviews were carried out with the managers of the facilities (team-leaders of community staffs and presidents and CEOs of the companies) that resulted in the collection of data from 13 different residential facilities. These facilities included the following: two high-intensity therapeutic communities [Type A communities in terms of the legislation of the Piedmont region (Com.1, Com.2)], two middle-intensity therapeutic communities [Type B communities in terms of the legislation of the Piedmont region (Com.3, Com.4)], a housing community (Com.5), and eight groups housed in apartments (AG 1, AG 2, AG 3, AG 4, AG 5, AG 6, AG 7, AG 8). Overall, 121 patients lived in the communities that were examined (male: 79.3%; age: 39.77±12.48). Among these patients, 42.5% suffered from schizophrenia or other psychotic disorders, 10% had diagnoses of personality disorders, 1.7% had mood disorders, 38.3% had diagnoses of both psychiatric disorders and substance abuse, and 7.5% had diagnoses both of psychiatric disorders and intellectual deficits. The patients had been living in the same residential facilities for 28.70±35.71 months.

Procedure

This research follows the methodological approach known as the mixed-method approach (Johnson & Owuebugzie, 2004). This method is characterised by a highly pragmatic understanding of research results that employs both qualitative and quantitative methods to answer research questions. It was considered useful to integrate qualitative and quantitative data to describe the different types of community residential facilities to achieve a flexible, yet logical, comparison of the different aspects of each facility. Specifically, this research involved the integration of the following diverse data sources: documentary material collected in the various residential facilities for psychiatric patients (e.g., from the website of the provider, the service charter, the project structure, the procedural rules and the therapeutic contract); semi-structured interviews with the management of each facility (to analyse cultural and organisational characteristics of the different residential facilities); and tabulated quantitative data about the human resources of each community residential facility. Table 1 shows a summary of the themes studied and the integration of the main data sources that were used to gather information in each area. Moreover, to identify differences between users of the different types of residential facilities studied, tabulated quantitative data about patients were collected. Specifically, the team leaders of the community staffs provided data about the socio-demographic (e.g., age and sex), clinical (e.g., diagnoses and age of diagnoses) and therapeutic statuses of each patient.

All interviews were recorded and transcribed and, together with documentary material, were analysed with a content analysis procedure. This procedure involved reading of the collected qualitative material by each member of the research group and the creation of an analysis matrix that defined codes that analytically described the emergent content of the interviews. Table 2 shows the codes used and their definitions.

This process was followed by insertion of the material collected (from interviews or documents taken from the different facilities) in a matrix of codes. Finally, the research group created a range of second-order categories, which began with first codes, through a process of group discussion. Based on these secondary order categories, our research group developed typologies of facilities studied that were used in the next phases of this study.

The quantitative data were analysed with SPSS 18 statistical software to highlight, using the simplest statistical indices (ANOVA's, Student's t-tests, χ^2), any differences or similarities amongst the diverse types of facilities identified and their users.

Table 1 – Summary of the themes studied in this research and the main data sources of this research

Themes investigated	Sources
Mission and goals of facility, as well as theoretical raison-d'être and development plan	Interviews with management, project structure and card services
General characteristics of the facility (name, date of establishment, type of facility by regional classification nomenclature, m^2, number of rooms, number of beds, hours of coverage, number of guests present, number of admissions and discharges in 2011)	Data collection sheet, interviews with management
The user route within the system (selection framework of the referrer, admission, evaluation, intervention, follow-up) (Ovretveit 1996)	Interviews with management
Activities and services offered by the facility (type of activity, leader, frequency of participation)	Data collection sheet, interviews with management
Staff organisation within the facility, staff management (characteristics, roles, recruitment process)	Data collection sheet, interview with management
Referrers' network	Interview with management

Table 2 – Analysis matrix for qualitative material

Content categories	Description
Mission and objectives of the facility	All the contents that refer to the culture and values that form the basis of the facility are reported here; to these are added objectives as they emerge from the collected material
Environmental characteristics of the facility	All physical, spatial and geographical location references to the facilities are reported here
User pathway	The general characteristics of the user pathway (but only limited information about the specific characteristics of each phase) are reported here
Selection, admission and assessment	This category refers to specific procedures carried out by the facility to select, evaluate and welcome guests
Treatment	Types of interventions provided by the facility, together with assumptions and values that underpin the different types of treatment proposed
Relationship with family members	Material relating to activities with and for the guests' families is reported here
Relations with the community	Material about the use of the surrounding community as a resource for the guest, including where the use of community resources to treat or rehabilitate guests are highlighted
Guest features	Material covering all attempts to categorise guests by diagnosis, and, more widely, by characteristics that would favour or not favour a successful outcome
Staff and their characteristics	Material defining or specifying the staff who are necessary and fundamental to the work in the facility is reported here
Staff management	Staff selection methods, the degree to which staff "sign-up" to the aims and ideals of the organisation, whether training is provided
Referral networks	Material describing referrers: their number, geographic proximity, and similarities or dissimilarities in terms of operational methodology

RESULTS

The content analysis we conducted permitted us to postulate that different types of facilities exist for psychiatric patients. These types were constructed based on two reference axes. The first axis was substantially related to the physical and environmental characteristics of these facilities and the support hours provided by the staff. We considered the number of beds for each structure, the space dedicated to each guest (in m3 for guest), the space for rehabilitative activity, and the number of hours dedicated by every member of the staff (psychiatrist, nurses, psychologist, educators). On this first axis, the first type of facility refers to large structures with a large number of beds that are normally defined as communities. In contrast, we also identified types of facilities that were akin to normal living spaces, i.e., smaller structures with a more limited number of beds, which are referred to here as protected homes and apartment groups and are grouped by the amount of support hours provided by the staff. The first group entailed very high support (24 hours), and the second group entailed more limited support (4, 6, or 8 hours per day).

The second axis that emerged from the other identified content categories refers to the operating methodologies chosen by each facility to work with guests. Here, we classified the structures in relation to their philosophy and their position about rehabilitation and recovery of patients with mental disorders. In this work we use, as theoretical framework, the model of recovery, in which the word "recovery" refers "both to internal conditions— the attitudes, experiences, and processes of change of individuals who are recovering—and external conditions—the circumstances, events, policies, and practices that may facilitate recovery. Together, internal and external conditions produce the process called recovery. These conditions have a reciprocal effect, and the process of recovery, once realized, can itself become a factor that further transforms both internal and external conditions" (Jacobson & Greenley, 2001, p. 482). In this axis, the first type of structures uses methodologies that focus on the care and management of what we call the guests' social skills. These facilities focus on rehabilitation, have a close relationship with the local community, and develop occupational therapy-based activities for their guests in partnership with local authorities and churches. In addition, the facilities that favour this approach were quite well organised and regulated and ensured that guests had supervised, secure accommodations. The second type refers to the operating methods of facilities that were strongly rooted in everyday life. These facilities favoured maintenance-stabilisation rather than change in their users. These facilities offered long-term residence spanning several years, the creation of space and autonomous projects that were "protected" by a low-key staff presence. These structures were particularly flexible, staff carried out educational activities and treatment, and planning was strongly determined on a case-by-case basis. The interviews revealed that these facilities had older guests with longer histories of illness.

Finally, there was a third type of facility that was characterised by operating methods that focused mainly on the mental and intrapsychic functioning of the guests. Establishments favouring this type of work were characterised by strong psychotherapeutic frameworks, which were typified by acceptance procedures, evaluations of the guests and well-defined activities that often occurred in groups. These facilities provided stays of finite durations (one or two years), and the goal of these facilities was to change the individual. From the perspective of these facilities, crisis-management is preferred to crisis-avoidance as a specific working objective. These facilities contained higher proportions of staff who were trained specifically for psychiatric and psychotherapeutic work (psychotherapists, psychiatrists).

Table 3 graphically represents these different axes and the different structures that were analysed in our research. In summary, we identified the following 8 types of residential facilities based on to their environmental characteristics and their care focus: apartment groups focused on social skills; protected homes focused on social skills; apartment groups focused on daily life; protected homes focused on daily life; community homes focused on daily life; apartment groups focused on mental and intrapsychic life; protected homes focused on mental and intrapsychic life; and community homes focused on mental and intrapsychic life. In addition, we hypothesise that another type of residential facility exists: communities focused on social skills; however, we did not identify this type of facility in the current study.

We tried to determine whether there were significant differences in the characteristics of the guests across the different types of facilities. Table 4 shows the results of this comparison. Specifically, the apartment-based groups focused on social skills, had guests who were young, had shorter residencies (approximately 1 year), and had projects that focused on employment. The protected homes focused on social skills, had guests with long periods of residency and were characterised by patients with lower educational achievement and psychotic disorders. These facilities offered socialisation activities such as sports and occupational therapy. Additionally, the guests of these facilities were characterised by histories of chronic psychiatric pathologies that most likely impacted their basic social skills. The groups in apartments and communities focused on daily life had guests who were generally older than the residents of the other types of facilities; the residents of these facilities also exhibited longer histories of mental illness. Community and protected homes that were focused on daily life tended to have rather long durations of guest residency. This finding did not apply to apartment groups; however, these data may not be applicable because apartment groups are newly created facilities. Facilities that focused on intrapsychic functioning (i.e., the community, protected home and apartment groups) tended to have shorter residency durations, lower average guest ages and shorter illness histories. The main guest activities were psychological and involved different types of group participation rather than individual participation. In addition to these activities, in the protected homes and apartment groups, there were significant introductory activities focused on job placement or self-management of leisure time.

Table 3 – Spectrum of facility types based on the environmental characteristics axis and the facility operational focus axis

AXIS OF ENVIRONMENTAL CHARACTERISTICS	AXIS OF OPERATIONAL FOCUS		
	Operation focused on social skills	**Operation focused on daily life**	**Operation focused on mental and intrapsychic life**
Apartment Group up to 5-7 guests, a few hours of daily support	*AG 3; AG 8* facilities that focus on the acquisition of social skills necessary for independent living. Strong focus on employment	*AG 7* facilities with low levels of support, focused on providing guests with the necessary support for independent living	*Com. 5* low-support facilities whose work focuses on the establishment and verification of independent living and crisis management abilities
Protected Home up to 5-7 guests, 24 h support / day	*AG 1* facilities involved in the acquisition of basic social skills needed for communal living and contact with the outside world	*AG 2; AG 5; AG 6* facilities devoted to the development of autonomy, with continuous work on the skills necessary for the self-management of daily routine	*AG 4* facilities devoted to the verification of the ability to self-manage intrapsychic stability, in the context of greater autonomy
Community 20-22 guests, 24 h support / day		*Com. 4* facilities devoted to the strengthening of autonomy for the management of everyday life amongst guests with high support needs	*Com. 1; Com. 2; Com. 3* facilities with a strong psychotherapy regime and an emphasis on change to help the guest better manage crisis situations

DISCUSSION

The literature reports a considerable diversity of models related to supportive housing and other facilities for psychiatric patients and the importance of research that seeks to identify specific features that discriminate among different settings (Fakhoury et al., 2002). The results of our study highlight different types of residential facilities for psychiatric patients that were defined initially by the characteristics of the structures and the aims and techniques of treatment, care and rehabilitation provided to users. Moreover, analyses of the characteristics of the patients who lived in different types of residential facilities revealed that different types of residential facilities matched different types of users. Specifically, our study identified different facilities based on their environmental characteristics and operative approaches to the care, assistance and rehabilitation of patients. The characteristics that we used to define the different type of facilities are not new; the literature related to the study of the quality of institutional care focuses on similar aspects, such as living conditions, the characteristics of interventions, and therapeutic relationships, among others. However, the studies comprising this literature often aim to identify the "ideal institution" (Taylor et al, 2009). In contrast, in this study, we have shown the characteristics of the structures that seem to be more suitable for certain patients depending, for example, on their ages and histories of psychiatric illness. The creation of types of psychiatric residential facilities may have the following important consequences: a) the specification, of guests better suited to a particular facility will improve the efficacy of treatment; b) assistance guidance in the development of health policies, particularly the planning of facility types based on user characteristics, the staff numbers required and the main activities to be delivered; and c) a description of some diagnostic and therapeutic methods based on current practice in different types of establishments that are recognized as effective in the national and international literature. In addition, identification of the

Table 4 – Differences between guest characteristics in different types of facilities

	Type of facility	Number of guests	Mean	Standard Deviation	F	p
Age (years)	AG social skills	4	33.00	2.70	4.88	<.001
	AG daily routine	5	55.60	6.54		
	AG mental functioning	7	35.86	5.24		
	PH social skills	5	48.00	7.90		
	PH daily routine	15	43.47	11.47		
	PH mental functioning	5	34.40	12.81		
	Com daily routine	20	47.25	15.04		
	Com mental functioning	60	35.70	10.84		
Time in facility (number of months)	AG social skills	4	13.75	9.60	6.99	<.001
	AG daily routine	5	3.00	.00		
	AG mental functioning	7	8.00	7.43		
	PH social skills	5	63.80	16.10		
	PH daily routine	15	27.73	29.85		
	PH mental functioning	5	11.60	9.65		
	Com daily routine	20	65.80	57.37		
	Com mental functioning	60	20.63	22.45		
Hospital admissions after patients are admitted to the facility	AG social skills	4	.25	.50	1.18	n.s.
	AG daily routine	5	.20	.44		
	AG mental functioning	7	.00	.00		
	AG social skills	5	.80	.83		
	PH daily routine	15	.67	1.29		
	PH mental functioning	5	.00	.00		
	Com daily routine	20	.30	.73		
	Com mental functioning	59	1.14	2.12		
Number of typical antipsychotics per guest	AG social skills	4	.50	.57	2.51	.019
	AG daily routine	5	1.00	.00		
	AG mental functioning	7	.14	.37		
	PH social skills	5	.60	.54		
	PH daily routine	15	.73	.59		
	PH mental functioning	5	.80	.44		
	Com daily routine	20	1.15	.81		
	Com mental functioning	59	.64	.63		

continued on next page

continued

	Type of facility	Number of guests	Mean	Standard Deviation	F	p
Number of atypical antipsychotics per guest	AG social skills	4	.50	.57	1.21	n.s.
	AG daily routine	5	.60	.54		
	AG mental functioning	7	.71	.75		
	PH social skills	5	1.20	.44		
	PH daily routine	15	.87	.64		
	PH mental functioning	5	.40	.54		
	Com daily routine	20	.70	.65		
	Com mental functioning	59	.59	.52		
Age at original diagnosis	AG social skills	4	17.00	3.46	2.36	.028
	AG daily routine	5	29.40	16.33		
	AG mental functioning	7	16.57	4.82		
	PH social skills	5	16.60	4.82		
	PH daily routine	14	26.86	12.45		
	PH mental functioning	5	21.20	12.21		
	Com daily routine	20	17.85	8.03		
	Com mental functioning	55	22.00	8.21		
Disease duration (years)	AG social skills	4	16.00	5.22	8.44	<.001
	AG daily routine	5	26.20	11.43		
	AG mental functioning	7	19.28	7.13		
	PH social skills	5	31.40	7.26		
	PH daily routine	14	15.07	9.14		
	PH mental functioning	5	13.20	6.30		
	Com daily routine	20	29.40	15.55		
	Com mental functioning	55	13.52	6.71		

Note. AG = Apartment Group; PH = Protected Home; Com = Community.

types of facilities that correspond to specific permutations of environmental characteristics and operating methodologies should not only allow for better choices of facilities based on how well suited those facilities are to each individual but also allow for the recruitment of appropriate staff for the care, treatment and rehabilitation culture of each particular facility. The development of facilities with appropriate work cultures and management climates would, by extension, also promote staff wellbeing.

Limitations

This study was conducted in a relatively small number of facilities and in a limited geographic area in Italy. Therefore, our results, although interesting, require further testing in a wider range of facilities that are not restricted to Italy. In this study, we used the same group of facilities to define the typologies and to verify the differences between the users who lived in these facilities. We believe that the characteristics of the facilities drove

the choices of patients and not vice versa; however, based on our data, it is not possible to reach this conclusion. We can only state that there was a correspondence between facilities with certain characteristics and guests with certain characteristics. Therefore, future research will be necessary to determine whether different types of residential facilities are appropriate for psychiatric patients with different characteristics.

CONCLUSIONS

Currently, a proportion of people with mental health conditions live in residential facilities. In the present research, we identified the following types of facilities based on their environmental characteristics and their focuses of care: apartment groups focused on social skills; protected homes focused on social skills; apartment groups focused on daily life; protected homes focused on daily life; communities focused on daily life; apartment groups focused on mental and intrapsychic life; protected homes focused on mental and intrapsychic life; and communities focused on mental and intrapsychic life. According to our data, the patients of these different types of facilities were also different in terms of age, diagnosis, and duration of disease. We believe that defining the types of residential facilities for psychiatric patients is important both for improving the definitions of the different approaches of care that are used in community interventions and for better defining the "quality of care" in each types of facilities.

Acknowledgements
This study was supported by FENASCOOP (Italian Association of Psychotherapeutic Community).

References

BRUNT, D. & HANSSON, L. (2002). A comparison of psychosocial environment of two types of residences for persons with severe mental illness: Small congregate community residences and psychiatric inpatient settings. *International Journal of Social Psychiatry*, *48*, 243-252.

DE GIROLAMO, G., BASSI, M., NERI, G., RUGGERI, M., SANTONE, G. & PICARDI, A. (2007). The current state of mental health care in Italy: Problems, perspectives, and lessons to learn. *European Archives of Psychiatry and Clinical Neuroscience*, *257*, 83-91.

FAKHOURY, W.K.H., MURRAY, A., SHEPHERD, G. & PRIEBE, S. (2002). Research in supported housing. *Social Psychiatry and Psychiatric Epidemiology*, *37*, 301-315.

JACOBSON, N. & GREENLEY, D. (2001). What is recovery? A conceptual model and explication. *Psychiatric Service*, *52*, 482-485.

JOHNSON, R.B. & OWUEBUGZIE, A.J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, *33*, 14.

KILLASPY, H., WHITE, S., WRIGHT, C., TAYLOR, T.L., TURTON, P., SCHUTZWOHL, M., SCHUSTER, M., CERVILLA, J.A., BRANGIER, P., RABOCH, J., KALISOVÀ, L., ONCHEV, G., ALEXIEV, S., MEZZINA, R., RIDENTE, P., WIERSMA, D., VISSER, E., KIEJNA, A., ADAMOWSKI, T., PLOUMPIDIS, D., GONIDAKIS, F., CALDAS-DE-ALMEIDA, J., CARDOSO, G. & KING, M.B. (2011). The development of the Quality Indicator for Rehabilitative Care (QuIRC): A measure of best practices for facilities for people with longer term mental health problems. *BMC Psychiatry, 11*, 35.

LORA, A. (2009). An overview of the mental health in Italy. *Annali dell'Istituto Superiore di Sanità, 1*, 5-16.

MORRIS, J., LORA, A., McBAIN, R. & SAXENA, S. (2012). Global mental health resources and services: A WHO survey of 184 country. *Public Health Review*, *34*, 1-19.

OVRETVEIT, J. (1996). *La qualità del servizio sanitario*. Napoli: EdiSES.

PICARDI, A., DE GIROLAMO, G. & MOROSINI, P. (2003). Le strutture residenziali psichiatriche in Italia. Risultati preliminare dalla fase 2 del progetto PROGRES. *Notiziario dell'Istituto Superiore di Sanità*, *2*, 3-9.

SANTONE, G., DE GIROLAMO, G., FALLOON, I., FIORITTI, A., MICCIOLO, R., PICCARDI, A. & ZANALDA, E. (2005). The process of care in residential facilities. *Social Psychiatry and Psychiatric Epidemiology*, *40*, 540-550.

TAYLOR, T.L., KILLASPY, H., WRIGHT, C., TURTON, P., WHITE, S., KALLERT, T.W., SCHUSTER, M., CERVILLA, J.A., BRANGIER, P., RABOCH, J., KALISOVÀ, L., ONCHEV, G., DIMITROV, H., MEZZINA, R., WOLF, K., WIERSMA, D., VISSER, E., KIEJNA, A., PIOTROWSKI, P., PLOUMPIDIS, D., GONIDAKIS, F., CALDAS-DE-ALMEIDA, J., CARDOSO, G. & KING, M.B. (2009). A systematic review of international published literature relating to quality of institutional care for people with longer term mental health problems. *BMC Psychiatry*, *9*, 55.

THORNICROFT, G. & TANSELLA, M. (2004). Components of a modern mental health service: a pragmatic balance of community and hospital care: overview of systematic evidence. *The British Journal of Psychiatry*, *185*, 283-290.

The evolution of the reading profile in children with developmental dyslexia in a regular ortographies

Maria Zonno[1], Maristella Scorza[2], Isabella Morlini[3], Giacomo Stella[2]

[1] Freelance Psychologist
[2] Department of Education and Human Science, University of Modena and Reggio Emilia
[3] Department of Economics, University of Modena and Reggio Emilia

ABSTRACT. Le ricerche dimostrano che la dislessia nelle lingue ad ortografia trasparente evolve, nel corso del tempo, in maniera differente per i parametri rapidità e accuratezza, in favore di quest'ultima. Scopo dello studio è quello di esaminare l'evoluzione dell'abilità di lettura lungo l'arco della scolarità obbligatoria, dalla scuola primaria a quella secondaria di secondo grado, nell'ortografia italiana. Inoltre, ha l'obiettivo di verificare se esistono differenti traiettorie evolutive in relazione alla severità del disturbo di lettura. Lo studio è stato condotto su un campione di 71 bambini dislessici italiani, secondo i criteri diagnostici stabiliti dal manuale diagnostico ICD-10 e dalla Consensus Conference. Il campione è stato suddiviso in due gruppi: un gruppo di dislessici lievi (n=36) e un gruppo di confronto di dislessici medio-severi (n=35). Tutti i partecipanti sono stati valutati almeno due volte in due diversi livelli scolastici. I confronti sono stati effettuati sulle prestazioni medie in ogni grado scolastico. I risultati rivelano traiettorie evolutive della capacità di lettura simili nel corso del tempo, in favore del gruppo dei dislessici lievi. I dislessici medio-severi mostrano un andamento che non è lineare nel corso del tempo, con un peggioramento nel corso dell'ultimo anno scolastico analizzato (1a secondaria di secondo grado), mentre il gruppo dei lievi mostrano un incremento costante nel tempo. Per quanto riguarda il parametro rapidità, entrambi i gruppi mostrano un incremento maggiore nella lettura delle parole e del testo, mentre rivelano un incremento minore nella decodifica delle non parole.

SUMMARY. Several researchers have demonstrated that dyslexia develops differently in shallow orthographies in terms of accuracy and speed. In fact, slow reading speed persists and accuracy improves. The aim of this study is to investigate the evolution of the specific reading disorder over the years of compulsory education, from primary to upper secondary school. Furthermore, it has the aim to verify if there are different evolutionary trajectories of reading skills in relation to the severity of the disorder. The study was carried out on 71 Italian dyslexic children, according to the diagnostic criteria established by the diagnostic manual ICD – 10 and the Consensus Conference. Two groups were selected: children who met criteria for mild dyslexia (mild dyslexics, with n=36) and a comparison group of moderate-severe dyslexics (n=35). All participants were tested at least twice in two different school grades. Comparisons were made on the average performances in each school grade. The results reveal similar patterns of growth over time in reading ability, with the mild dyslexics group outperforming the moderate-severe dyslexics group. The performance trajectory for the moderate-severe dyslexics shows some plateaus and a decrease in performances in the last year analyzed (1st upper secondary school) while the trajectory for the mild dyslexics always show increases in performances. All subjects show a steady increase in word and text reading speed and a slower improvement in pseudo-word decoding.

Keywords: *Developmental dyslexia, Reading, Regular orthographies*

INTRODUCTION

Most of the experts agree that dyslexia is a lifelong condition that can spontaneously improve and change in form (Tressoldi, Stella & Faggella, 2001). Several researchers agree that in consistent languages (characterized by high grapheme-phoneme correspondence) the critical sign of dyslexia concerns the speed in decoding ("speed dyslexia", Wimmer, 1993).

The slowness in reading persists, especially in the reading of pseudowords where there is a lower increase in speed that seems to reach a ceiling ("ceiling effect") at the end of the secondary school level; in the reading of the text and of the words, instead, it occurs the lexical effect (Shaywitz et al., 1999; Stella, Savelli, Scorza & Morlini, 2010; Tressoldi et al., 2001).

As regards instead the parameter accuracy, several authors show that the time lead to an increase in the accuracy of the master such that the gap between dyslexics and typical readers tends to shrink; we also know that in the transparent languages there is a lower number of errors compared to opaque languages (Holopainen, Ahonen & Lyytinen, 2001; Jimenez, 2012; Paulesu et. al., 2001; Tressoldi et al., 2001).

The longitudinal study seems to be a valid tool to explain the evolution of dyslexia. In fact, such analyses identify the parameters that remain unvaried during the developmental phases, recognizing the predictive signs of the severity and the persistence of the disorder. These investigations are also useful to establish suitable rehabilitation plans for dyslexics. The Connecticut Longitudinal Study (Shaywitz et al., 1999) is one of the first perspectives on the evolution of the disorder. The in-depth exploratory and follow-up study of dyslexic children into adulthood is carried out on a sample of 445 children. The results of the study show that slow reading speed and phonological deficits persist during adolescence and adulthood, whereas decoding accuracy improves.

Other longitudinal studies aim to understand why some children are vulnerable to the acquisition of reading skills, such as the the Jyväskylä Longitudinal Study of Dyslexia (Lyytinen et al., 2006) and the Dutch Dyslexia Programme (Van der Leij et al., 2013).

The Jyväskylä Longitudinal Study followed 200 Finnish children from birth to school age. Half of these children had a family history of reading problems and were considered at risk for dyslexia; the other half were not at risk. They have identified four subgroups with differential trajectories to early reading. The results revealed that there are at least three troubled routes along which a child may ultimately encounter difficulties in reading acquisition. The most explicit routes are characterized by problems in either phonological awareness, naming speed, or letter knowledge problems that increase in severity with age (Lyytinen et al., 2006).

The Dutch Dyslexia instead analyzed a sample of 180 children with a familiar risk of dyslexia and a comparison group of 120 children without familiar risk of dyslexia and followed them from 2 months old up to 9 years. With regard to precursors of reading disability, the children were divided into three groups: familiar risk (FR) children with and without dyslexia, and controls. The results showed that regarding reading development, the FR children with dyslexia read less fluently since first grade onwards than the other two groups; the reading fluency of the FR children without dyslexia, instead, was at an intermediate level between the other groups and, furthermore, their word reading fluency gradually improved relative to the controls. By fifth grade, they had managed to catch up on word reading fluency, although they were still significantly slower than the controls on pseudowords reading fluency, indicating problems with word reading when sublexical orthographic knowledge is required (Van der Leij & Van Daal, 1999; Van der Leij et al., 2013).

The regular orthographic system of the Italian language makes it relatively easy to learn to read and write. In fact, reading and writing skills consolidate in the first two school years and children seem to be able to read 95% of a list of high-frequency words, at the end of primary education. (Scorza et al., 2015; Zoccolotti, De Luca, Di Filippo, Judica & Martelli, 2008). Tressoldi (1996) finds an average increase of .5 syllables per second during each year until the end of the lower secondary school, while the average text reading speed is 6 syllables per second.

Other works (Arina, Iervolino & Stella, 2013; Stella & Tintoni, 2007) show that decoding speed and accuracy still evolve after lower secondary school. There is a significant and persistent difference between dyslexics and normal readers in terms of decoding speed. Both groups improve their reading speed every year but variations in performance across grades can become more marked. Normal readers increase their reading speed by .5 syllables per second in both words and text reading and dyslexics by .3 syllables/second, less than their peers do (Tressoldi et al., 2001). In fact, the reading speed of dyslexic students attending the third year at lower secondary school is equal to that of normal readers in early literacy. The analyzed performances reveal that the lexical effect (Ziegler,

Perry, Ma-Wyatt, Ladner & Schulte-Korne, 2003; Zoccolotti & Burani, 2010) influence text reading and the words superiority effect can influence pseudowords reading.

However, many studies reveal that the progress is strictly related to the level of severity detected during infancy, as mild dyslexics improve more than severe ones (Lami, Palmieri, Solimando & Pizzoli, 2008).

The longitudinal study by Stella et al. (2010) is conducted on a sample of 35 dyslexic children. This study demonstrates that mild dyslexics (17 subjects) improve consistently in text and word reading in upper secondary school. However, their speed improvement is markedly slower in pseudowords decoding and they do not make any progress in more advanced education levels. It is possible to imagine a sort of "ceiling effect" on speed improvement when decoding new words, similarly to what happens to adult compensated dyslexics.

The group of severe dyslexics (17 subjects) shows a much lower increase in reading speed compared to the group of mild dyslexics.

In text and word-reading tests, the severe dyslexics in upper secondary school have a reading speed comparable to that of normal readers attending class 3 at primary school. In pseudo-word reading, they present further difficulties and they do not even reach the level of normal readers in class 2. Severe dyslexics increase to 1 syllable per second during the entire period of compulsory education (Stella et al., 2010).

In terms of accuracy, there are not substantial differences between mild and severe dyslexics. Both groups show notable improvements, which are very close to the normative values of the population (Lami et al., 2008; Stella et al., 2010). Campanini, Battafarano & Iozzino (2010), however, reach a different conclusion in their transversal study conducted on 291 dyslexic young subjects. They show, in fact, that the number of errors rises considerably in all classes and even tends to increase with education, leaving a widening gap between normal readers and dyslexics. Tucci, Savoia, Merella & Tressoldi (2013) replicate Stella's study (Tressoldi et al. 2001). They examine the natural evolution of reading acquisition in 57 dyslexic young subjects using a transversal-longitudinal study. The results show that there is still a gap between dyslexics and normal readers in terms of speed as school grades increase. Regarding accuracy, the number of errors tends to decrease in dyslexics but it is still quite high compared to that of their normal-reading peers, especially in the words and text-reading tasks.

Many authors agree that in a regular orthography like Italian, time produces an increase in accuracy control that reduces the differences between normal readers and dyslexics. In terms of decoding speed, there is instead a broad gap between both groups, despite a slight increase. These findings show that the critical parameter for dyslexia in regular orthographies is decoding speed. Hence, we can speak of speed dyslexia (Wimmer, 1993).

The Italian studies are consistent with those on the evolution of the disorder carried out in other countries with regular orthographies. Most of the international research on the developmental dyslexia, in fact, suggest that the reading difficulties encountered are mainly two, depending on the kind of orthography.in fact, phoneme-grapheme decoding accuracy in significantly low in opaque orthographies, whereas reading speed is slow in shallow orthographies (Scortichini; Gasperini, Scorza, Boni & Stella, 2015). For example, Wimmer & Mayringer (2001; 2002) conduct studies on German, Leinonen et al. (2001) and Holopainen et al. (2001) on Finnish, and Serrano & Defior (2008), Jimenez (2012) and Jimenez & Hernandez-Valle (2000) on Spanish. They show that children have problems both in speed and accuracy in pseudowords decoding. Undheim (2009) diagnoses a sample of Norwegians with dyslexia at the age of ten. Conducting a follow-up study of the same sample at 16-23 years old, he notices that all reading times are much higher than the normative values especially in pseudo-word decoding. Recently Caravolas (Caravolas, Lervag, Defior, Malkova & Hulme, 2013) has conducted a longitudinal study on reading acquisition in English, Spanish and Czech. The results show a slower development of reading abilities in English compared to other two orthographies that are more consistent.

Goswami and Ziegler (2005; 2006) explain the relationship between reading development and linguistic context. The Grain Size Theory demonstrates that there are substantial discrepancies between different spelling systems. In some languages such as English or Danish many different sounds correspond to a single grapheme, while, in orthographies like Italian or Spanish, a single grapheme corresponds to a single phoneme (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001; Scortichini et al., 2015).

Therefore children learning to read in orthographies considered opaque make more mistakes and are less fluent compared to children reading regular orthographies. The accuracy parameter refers to a cross-cultural study conducted on 36 dyslexics from France, England and Italy (12 for each

country) compared with a control group of 36 subjects equally distributed. All participants are administered both phonological short-term memory tasks and reading tests. In the short-term memory tests, the groups show a deficit, whereas the Italian dyslexic group achieve the best score in accuracy. The authors conclude that there is a universal neurocognitive base for dyslexia and that the orthographic structure of the languages rather than dyslexia causes disparities between the reading abilities (Paulesu et al., 2001).

In conclusion, slow reading speed seems to be the main problem in adolescence, whereas accuracy tends to improve with education. Subjects suffering from developmental dyslexia (DD) present a phonological deficit; this is why they read more slowly and less fluently than normal readers do. It is then essential for them to have more time to activate all the cognitive and linguistic (semantic-lexical) abilities, which compensate for the lack of decoding skills (Tucci & Tressoldi, 2009). The neuropsychological profile of adult dyslexics is particularly attractive because it explains the evolution of the disorder over time, as it affects other aspects of the cognitive function besides the reading difficulties. Kinsbourne (1991) carries out a study on 34 adults distributed in two groups: "severe" and "compensated" dyslexics. Severe dyslexics show deficits in verbal fluency, in rapid automatic naming, in verbal acquisition tests and temporal judgements. Compensated dyslexics perform poorly, instead, in rapid automatic naming and verbal fluency (Ghidoni, 2011). Hatcher, Snowling & Griffiths (2002) asses a sample of 23 dyslexics, whose average age is 25 years. The authors note that they perform poorly in pseudowords decoding, spelling, digit span, and writing speed. The personal experiences of the subjects reveal difficulties in manipulating data and organizing their work (Martino et al., 2011). Maughan et al. (2009) have conducted a significant follow-up study on a group of forty years old dyslexics, thirty years after the diagnosis of the disorder. The subjects still show persistent spelling deficits affecting the daily reading and writing activities (Ghidoni, 2011). Re, Tressoldi, Cornoldi & Lucangeli (2011) carry out a study on 104 university students from Padova. The results reveal that the average reading speed was four syllables per second, which was adequate for studying autonomously. However, old difficulties re-emerged under stress conditions (such as articulatory suppression) affecting the quality of the learning. Recently, Ciuffo et al. (2014) have conducted a study on silent reading, which is the standard reading form in teens, university students and adults. The results suggest that both normal readers and dyslexics improve their speed in silent reading rather than in loud reading. The improvement achieved by the dyslexic group, though, is clearly inferior to that of normal readers. It is plausible to suppose the presence of a structural deficit in automated reading, which is the process that promotes lexical access and facilitates the reading activity. These data confirm that there is a striking difference between dyslexics and normal readers in silent reading mode. In fact, dyslexics' top reading speed is 6.15 syllables/second, whereas the control group score 10.75 syllables/second. The results also emphasize the reduced speed difference between the loud reading test (4.89 syllables/second) and the silent reading test. The comparison reveals that there is a specific deficit in the recognition process, which is the basic structure of the decoding activity. This cognitive deficiency seems to be the cause of decoding issues rather than the verbal articulation of the written words required in laud reading.

AIMS AND SCOPE

The present study aims to investigate the evolution of the specific reading disorder over the years of compulsory education from primary to upper secondary school. Furthermore, it has the aim to verify if there are different evolutionary trajectories of reading skills in relation to the severity of the reading disorder.

The research examines a sample of subjects diagnosed with dyslexia between the second and the third class of primary education. All participants are re-assessed over the years of compulsory education at least twice and no more than seven times. Comparisons are made on the average performances in each school grade. A proportion of the sample in each school year is dropped from the subsequent year and replaced with different children. Therefore, each pair of samples coming from two different school grades have some children in common and some other children present in only one of the two samples. The study aims to analyze the development of reading abilities in dyslexics, through a series of reading tests, and to characterize and compare the pattern of grow over time in word, pseudowords and text reading. Two groups of participants are identified, according to the seriousness of the disorder (mild and moderate-severe), in order to examine the different evolution of the reading abilities.

PARTECIPANTS

The selected 71 participants are children enrolled in compulsory school, coming from different regions of Italy and diagnosed with dyslexia between the second and the third class of primary education. The medical diagnoses of the subjects comply with the diagnostic manual ICD-10 and the Consensus Conference (2007; 2011), in agreement with the discrepancy criterion between reading ability and general intelligence. All subjects were required to reach a performance QI and a Verbal QI >85[1] (Verbal QI score obtained in the PPVT-R, 2000) and were assessed at least twice in two different school years, to evaluate their reading disorder. Assessments have been done during the period 1998-2015. Children were recruited from patients consulting a private professional studio. Of the seventy-one participants, 47 are males and 24 females. This interesting detail is consistent with the hypothesis that dyslexia affects more males than females. In fact, the risk of developing ED is 2.5 times higher in males than females (Consensus Conference, 2011). Another remarkable aspect is that there are three couples of brothers of which two twins.

PROCEDURES AND TOOLS

All reading profiles are evaluated with the following tools:

- Words and pseudowords reading tests from the Battery for the evaluation of developmental dyslexia and dysorthography, (Sartori, Job & Tressoldi, 1995, 2007). Tests differ in features in each class and are adequate to the educational level of the child.
- MT reading tests for children in primary and lower secondary school (Cornoldi & Colpo, 1995, 2012) and MT advanced reading test (Cornoldi et al., 2010) for students in upper secondary education.

The reading ability is evaluated considering speed and accuracy. Speed is measured both with the overall reading time (in seconds) and the number of syllables per seconds read (fluency). For comparisons between these two measures and a comprehensive discussion about the problem of measuring reading speed in reading tests we refer to Cornoldi & Colpo (1995, 2012), Lorusso, Toraldo & Cattaneo (2006), Morlini, Stella & Scorza (2013, 2014, 2015).

Accuracy is measured with the number of errors made in all three tests.

All subjects are divided into the following two groups according to the reading performace in the first assessment:

- Mild ED group: if the reading time, in seconds, in the list of words, falls between the second and third standard deviation.
- Moderate-severe ED group: if the reading time, in seconds, in the list of words, falls between the second and third standard deviation.

METHOD

Comparisons are made on the average performances in each school grade. The study used a rotating sample design with participants interviewed at least twice during the years of compulsory education. A proportion of the sample in each school year is dropped from the subsequent year and replaced with different children. Therefore, each pair of samples coming from two different school grades have some children in common and some other children present in only one of the two samples.

RESULTS

Reading Development: comparison between dyslexics and control group

First, the average pattern of all participants affected with a reading disorder is compared to the normative values of the tests. The curve of grow in performances is similar in all the three tasks (words, pseudowords and text reading) and reveals a substantial gap between the dyslexics' decoding ability and that of their normal-reading peers. The gap increases as the level of educational attainment increases. The difference in performances between dyslexics and non-disabled children is greater in reading of the list of words. Figure 1 shows the decoding speed trend of the dyslexic subjects in words and pseudowords reading tasks in comparison with the control group. Figure 2 shows the decoding speed trend of the

1 The tests used to assess the cognitive abilities are: the Raven's Progressive Matrices (CPM 47 ; SPM 38) and scales WISC - III (for the subjects assessed before the year 2012) and WISC–IV.

Figure 1 – Words and pseudowords reading speed: comparison between collected data for dyslexics and control values

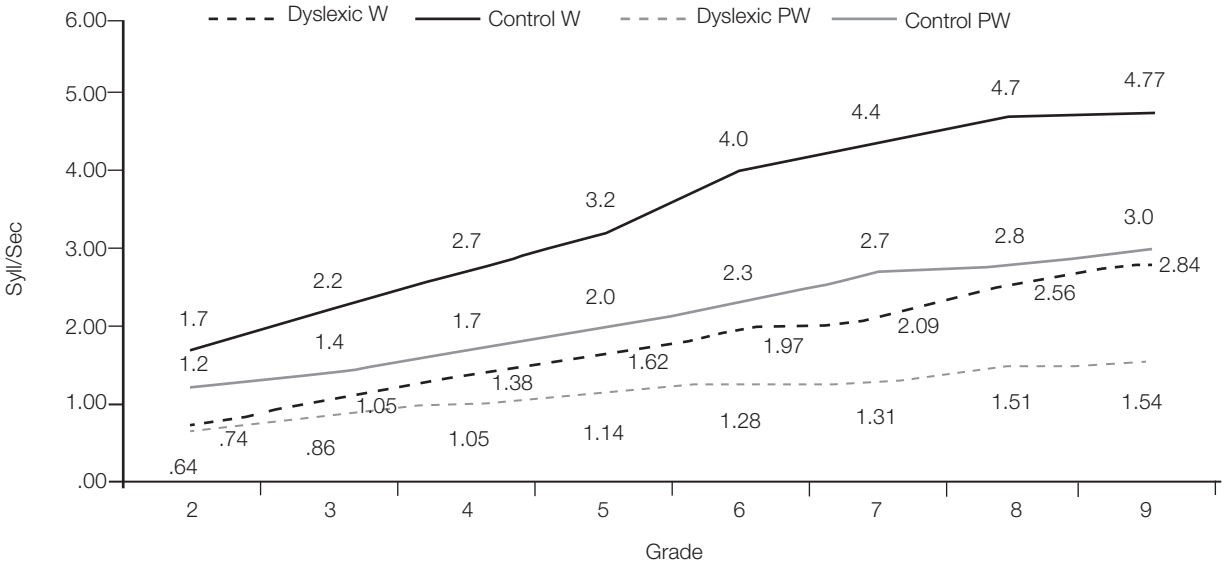
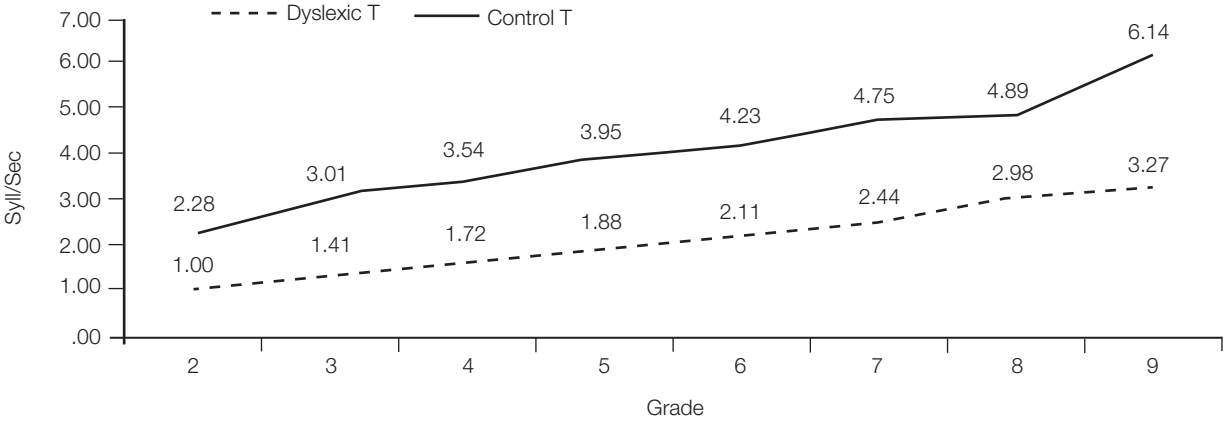


Figure 2 – Text reading speed: comparison between collected data for dyslexics and control values



dyslexic and non-dyslexics subjects in the text reading task. For the text reading task, the gap between disables and non-disables students tends to increase with education, as long as for the word and pseudowords reading tasks. Indeed, in the first year at upper secondary school, dyslexics read about 3.27 syllables per second, which are about half of the syllables read by a normal reader. The average rate of grow per year in reading the list of words is .29 syllables per seconds for dyslexics and .44 for the control group. The greatest increase occurs between class II and III of lower secondary school (.54 syllables/second) and the lower increase between class IV and V of primary school (.16 syllables/second). In reading the list of pseudowords the average rate of grow per year is .12 syllables per seconds for dyslexics and .27 for the control group. In reading the text the average rate of grow is .31 syllables per seconds for dyslexics and .48 for the control group. The speed in reading the list of words seems to best separate disable and nondisabled readers and to be the best predictor for dyslexia. The greater improvement in performances in words reading is probably due to the high frequency of the terms used. Decoding new words in the pseudo-word tests is obviously more challenging.

For what concern reading accuracy, the best performances and improvements are observed in the words reading test. In this test, dyslexics still keep improving theirs skill in the advanced educational stages, while in the other tests the curve of grow in the last few years have some plateaus or show increases in the number of errors. In pseudo-word reading, the improvement is less evident and the number of errors is quite high even in the advanced educational stages. The text-reading test reveals a nonlinear and non-monotonic trend over time. This trend confirms that text reading is the most difficult task for dyslexics. Figure 3 reports averages errors in words, pseudowords and text reading for dyslexics and the control values of the tests.

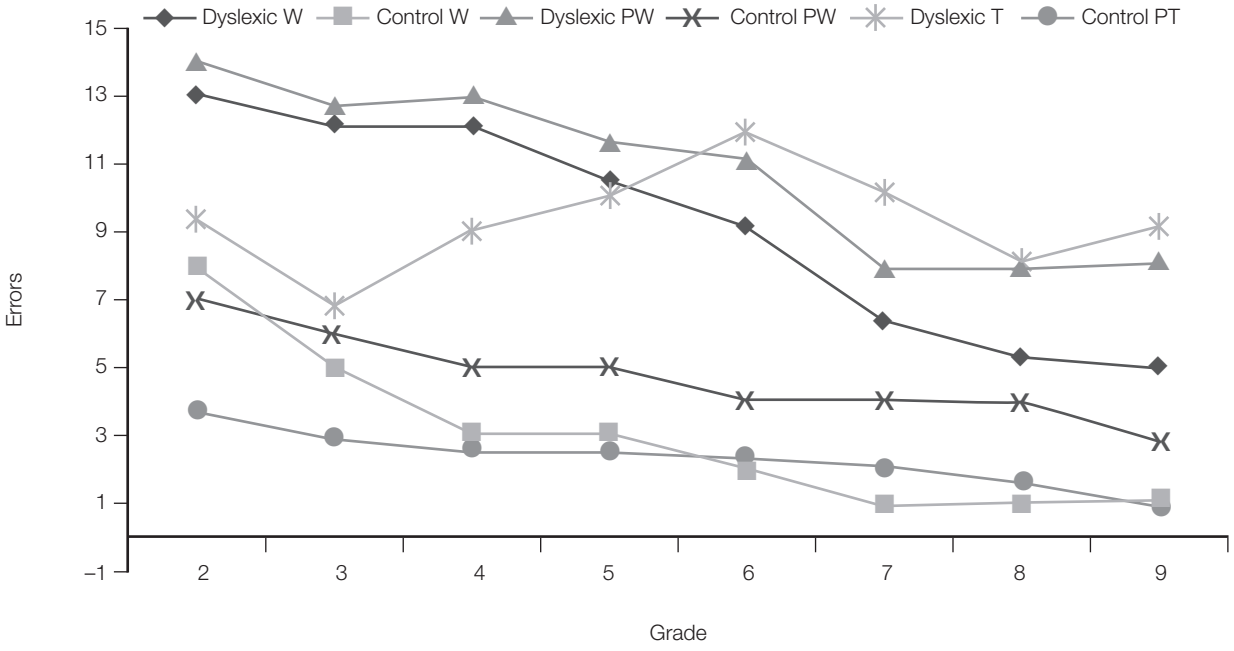
Analyzing averages values for speed and accuracy in each school year, we may draw some conclusions:

- Reading skill in dyslexics improves both in accuracy and in speed, during the eight analyzed years of compulsory education.
- In text reading, both the speed and the number of errors increase with education. This may be due to the fact that the difficulty in reading the text affects more the accuracy of reading (number of errors) than the speed (syllables per second read) and the higher the reading speed, the greater the number of errors. The result of this study is consistent with the available literature on the subject (Stella et al., 2010; Tressoldi, 1996; Tressoldi et al. 2001).
- Speed, in the word reading test, seems to be the most reliable predictive indicator of the future development of the reading ability. In word reading, the reading speed of dyslexic children increase by .29 syllable/seconds per year. In pseudo-word reading, the average improvement is .12 syllables/second per year. The greatest increase occurs between class II and III of primary school (.22 syllables/second) and between class II and III of lower secondary school (.20). A minimal improvement is shown between lower and upper secondary school (.03 syllables/second). This pattern confirms the "ceiling effect".

Reading Development: comparison between mild and severe dyslexics

Another aim of this research is the comparison of the dyslexics according to the severity of their diagnosis. In order to model changes in reading over the time span of the study, we have interpolated a linear regression line on the yearly average values for the groups of mild dyslexics and severe dyslexics and for all dyslexics (average curve). Because of the relatively small sample sizes of mild and severe dyslexics in each year, we have chosen a linear rather than a quadratic or nonparametric

Figure 3 – Errors in words, pseudowords and text reading: comparison between collected data for dyslexics and control values



function. The F-test in all regression lines, except for the line interpolating the number of errors in the text reading, show a very good fit: all *p*-values are smaller than .001 and indicate that the slope of the regression line is significantly different from zero, even for $\alpha = .001$ level of the test. The only pattern that cannot be interpolated with a straight line is the pattern relative to the number of errors in the text reading. Comparing the slopes of the curves, we see that in each measure (syllables per seconds, seconds and number of errors) and in each test (words, pseudowords and text) mild dyslexics demonstrate the highest level of improvement in reading performances, the average group demonstrate the next highest level and the severe dyslexic the lower level.

As regards the reading of the words, Figure 4, 5 and 6 show the development of speed and accuracy in the words reading test, over the period examined.

Regarding speed, performances of severe dyslexics still remain quite distant from the average in the last years. A severe grade 9 dyslexic reads 1.78 syllables per second and reaches the reading speed of a grade 2 normal reader (1.70

Figure 4 – Words reading speed for dyslexics: comparison between mild and moderate-severe dyslexics

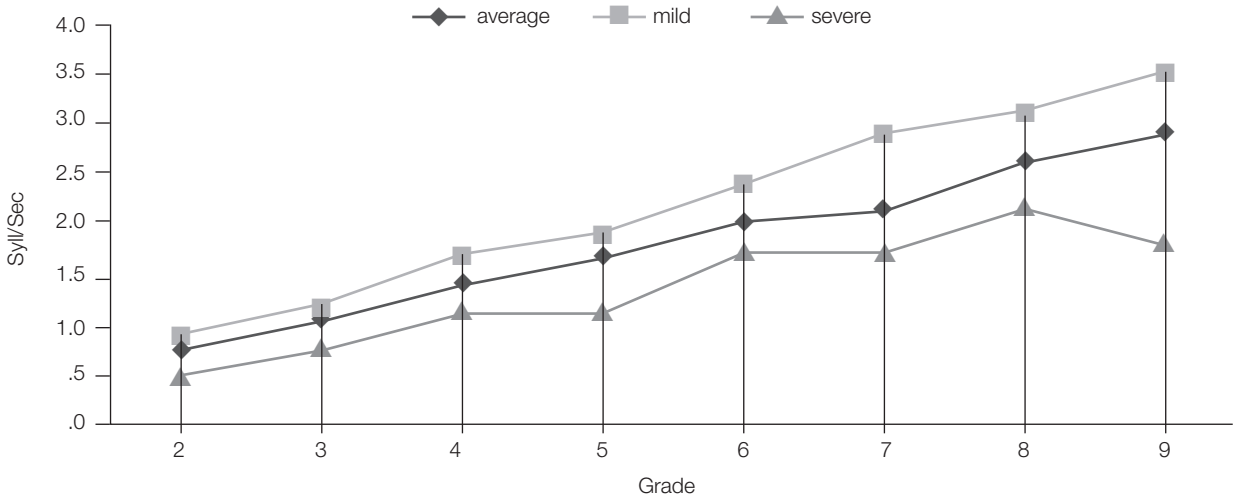


Figure 5 – Words reading fluency for dyslexics: comparison between mild and moderate-severe dyslexics

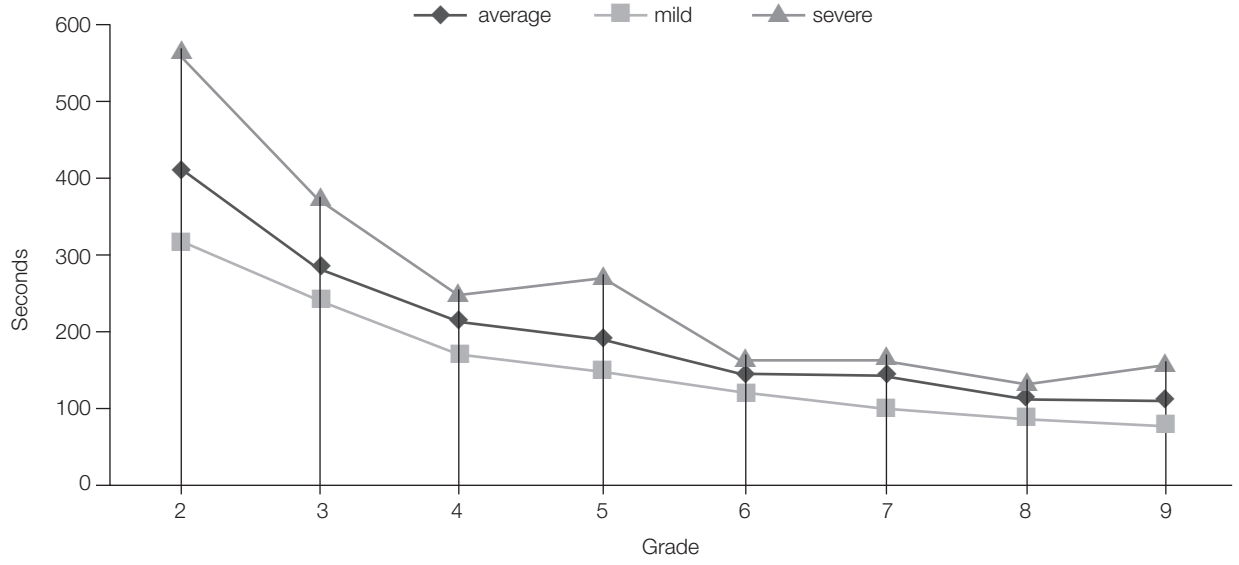
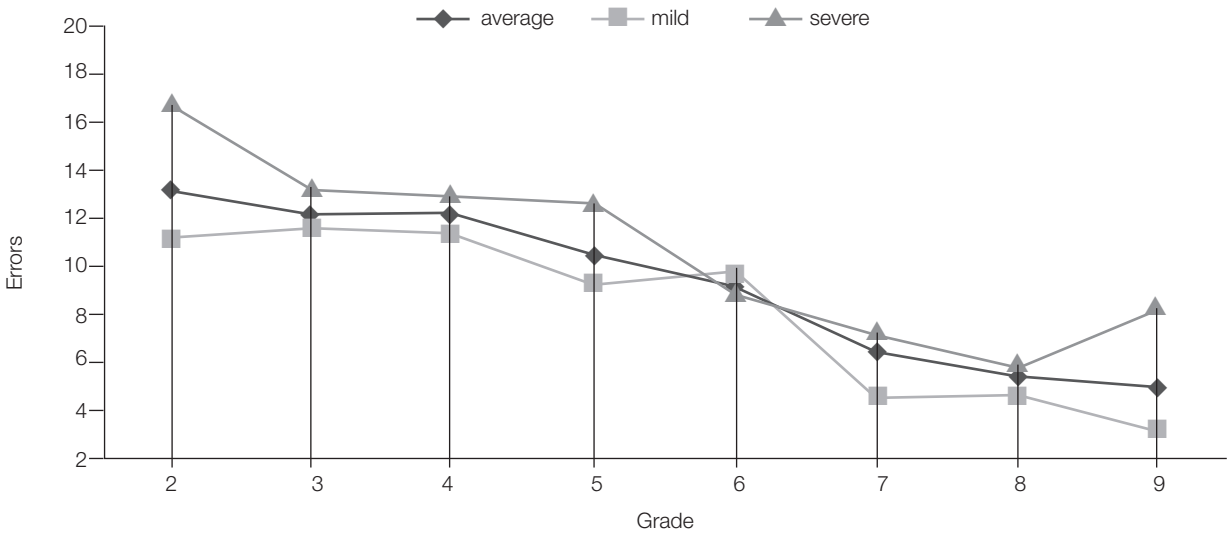


Figure 6 – Accuracy in reading words for dyslexics: comparison between mild and moderate-severe dyslexics



syllables per second). The group of mild dyslexics shows annuals increases of performances while severe dyslexics in some years show constant or even decreasing performances.

Considering accuracy, in all grades, except for grade 6, mild dyslexic students make fewer errors than the severe ones and their overall performance tend to improve more rapidly, especially in upper secondary school when the distance from the normal readers seems to decrease and the distance between mild and severe dyslexics reach the maximum value.

As regards the reading of the pseudowords, Figure 7, 8 and 9 show the development of speed and accuracy in the pseudowords reading test, over the period examined.

Severe dyslexics present a substantial distance from the average, confirming their phonological difficulties. Results show that a grade 9 student who reads about 1.14 syllables/seconds does not even reach the average speed of the grade 2 control group (1.20 syllables/seconds) after eight years of schooling.

In pseudowords reading speed, the curve of grow for mild dyslexics is smoother than the curve of grow for severe disables students. Severe disable students show greater differences in performances between years and also decreases in performances in some years. Regarding accuracy, the pattern is not monotone both for mild and severe dyslexics and the differences in performances are less evident.

Both groups make more mistakes in this test rather than in the words reading test. Mild dyslexics make a steady and remarkable improvement compared to the group of severe dyslexics, especially in the advanced educational years where they reduce the distance from the control group.

Finally, as regards the reading of the text, Figure 10 and 11 show the development of speed and accuracy in the text reading test, over the period examined.

Mild dyslexics improve their decoding speed by .39 syllables/second, whereas severe dyslexics by .23. This result confirms a minor improvement in the decoding ability that creates a considerable gap between dyslexics and normal readers (increase by .55 syllables /sec). The distance between performances of mild and severe dyslexics increases in the last school years.

Regarding accuracy, both groups show a nonlinear trend, making more mistakes between primary and lower secondary school. However, mild dyslexics make fewer mistakes than the severe ones do. The cause of the great number of errors is probably the increasing length and difficulty of the chosen texts. As long as for speed, the distance between performances of mild and severe dyslexics increases in the last school years.

Moreover, another interesting finding concerns the characteristics of the increase detected in the decoding speed of the words and the text in the dyslexic groups. Both groups show an almost identical increase in the speed of reading of the words and the text (.37 syllables/second in the words and .39 syllables/second in the text for mild dyslexics; .22 syllables/second in the words and .23 syllables/second in the text for the moderate-severe dyslexics) contrary to the typical readers that show a significantly higher average progress in the text reading compared to words reading (.55 syllables/

Figure 7 – Pseudowords reading speed for dyslexics: comparison between mild and moderate-severe dyslexics

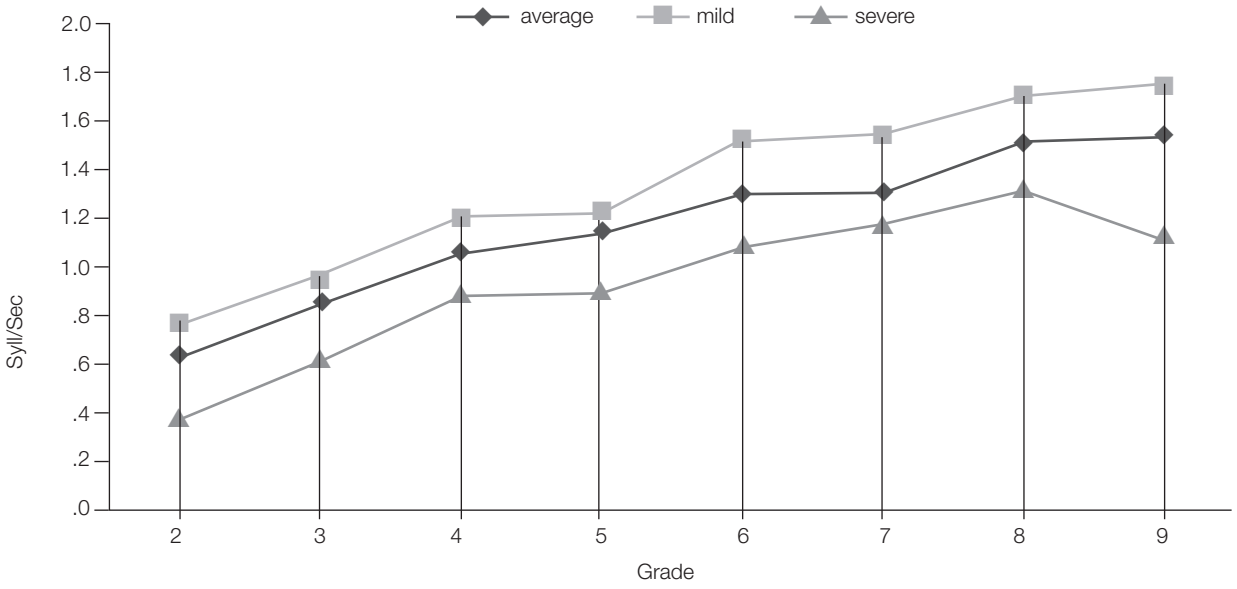
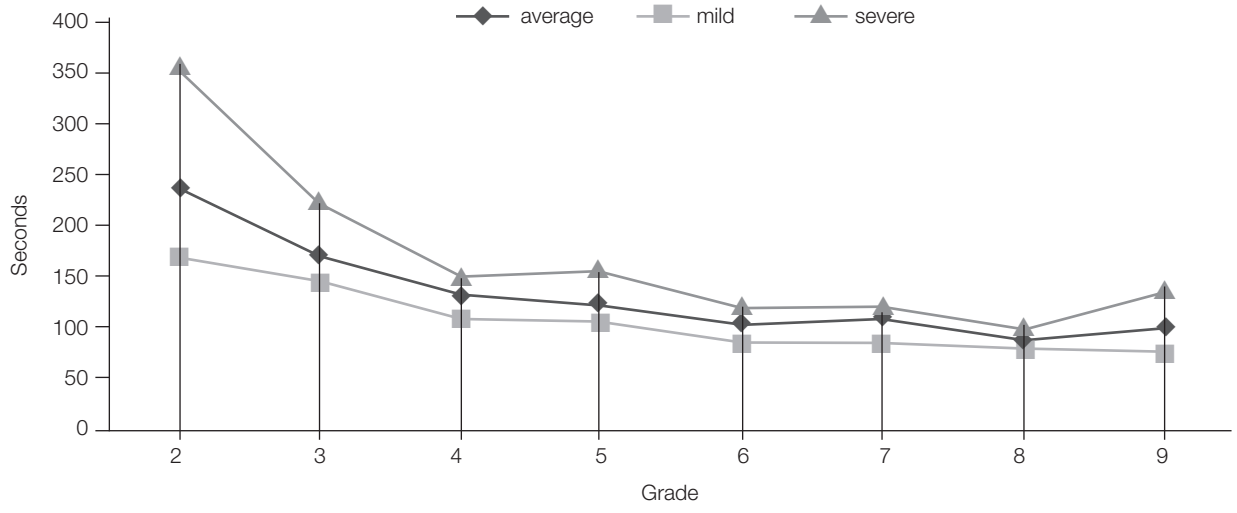


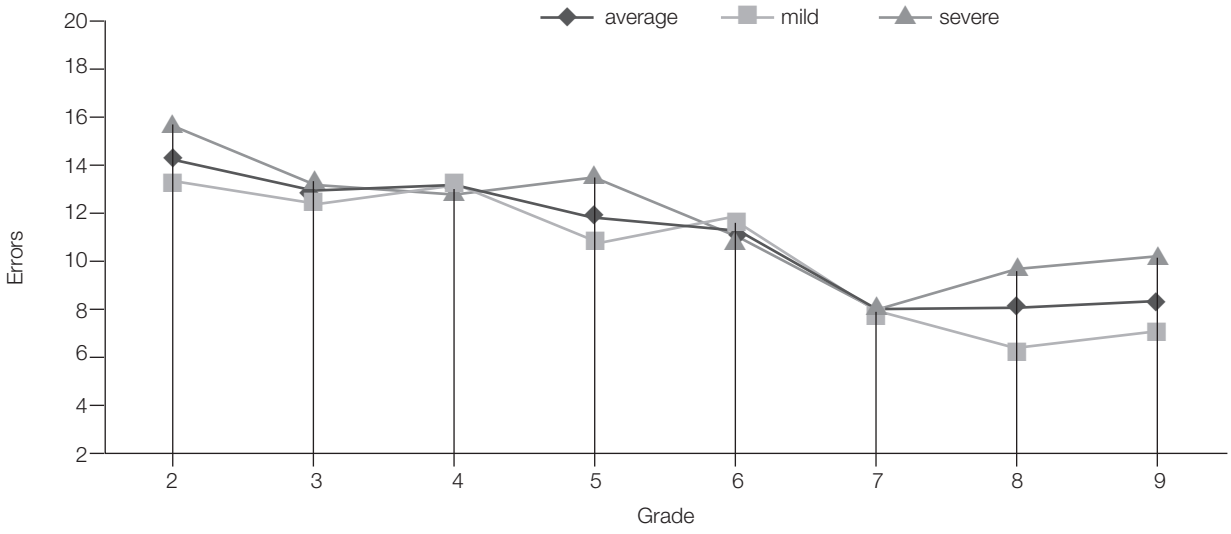
Figure 8 – Pseudoword reading fluency for dyslexics: comparison between mild and moderate-severe dyslexics



second in the text and .44 syllables/second in the words; Sartori et al., 1995, 2007).

The lack of advantage in the text reading than the words reading could be attributed to two different factors: inefficiency in some aspects of visual processing (crowding effect; Martelli, Di Filippo, Spinelli & Zoccolotti, 2009 - or preview effect – McCandliss, 2012) or weakness of some linguistic processes. In this last case, given that there is a difference in the speed increase between words and pseudowords, the inefficiency does not concerns lexical aspects, but it regards the facilitation that comes from the "linguistic knowledge" (Leonard, 2009). In fact, it would produce some advantages in lexical access deriving from implicit knowledge gain about utterance construction (Stella, 2013). With regards to the visual processing, there should be considered the advantages derived from the manipulation of text spacing (Zorzi et. al.,

Figure 9 – Accuracy in reading pseudowords for dyslexics: comparison between mild and moderate-severe dyslexics



2012), while the linguistic processes imply a revision of the role of lexical factors in reading process, by distinguishing lexical aspects (word recognition) from semantic-syntactic ones (related to the textual structure).

Table 1 reports the slope, the value of R^2 and the *p*-value for the F test in the regression lines fitted on the yearly average values for the groups of mild dyslexics and severe dyslexics and for all dyslexics (average regression line). The accuracy in reading the text is the only measurement that cannot be fitted by a linear line. All other measurements show a very good fit (all *p*-values are smaller than .01 and most of them are smaller than .001).

Regarding speed, the average yearly improvement (given by the slope of the regression line) of mild dyslexics is higher than the average improvement of moderate-severe dyslexics, in all tasks. Mild dyslexics have an annual improvement of .37 syllables per second in reading the words, .14 syllables per second in reading pseudowords and .39 syllables per second in reading the text. Moderate-severe dyslexics have an improvement of .22, .12 and .23 syllables per second, respectively. In words and text reading, both groups have an higher improvement than in pseudowords reading: this confirms the "ceiling effect" in decoding new words (Stella et al., 2010).

In reading fluency (measured with the time in seconds), severe dyslexics show a higher yearly improvement than mild dyslexics while in reading accuracy severe dyslexics improve better in words reading and mild dyslexics improve better in pseudowords reading.

DISCUSSION

Results of this study show that both the decoding speed and the decoding accuracy in dyslexics improves over the years of compulsory education. However, the gap between dyslexic and typical readers remains and that the decoding deficits recorded a different development in relation to the two parameters of speed and accuracy, in favor of the latter.

In shallow orthographies, Wimmer (1993) has noted how the reading disorder is much more evident in terms of speed and accuracy. Sometimes reading can be completely or almost accurate but is typically slow, with many pauses and hesitations. The data of our sample confirm the findings in international studies (Holopainen et al., 2001; Jimenez, 2012; Paulesu et. al., 2001; Shaywitz et al., 1999; Van der Leij et al., 2013) and other Italian studies (Stella et al., 2010; Tressoldi et al., 2001; Tucci et al., 2013): the gap between dyslexic and typical readers is progressively reduced for the parameter speed reading of words and text, while in the pseudo-word occurs less increase ("ceiling effect"). The accuracy improves instead to a greater extent in the words, while in the text and in the pseudowords, while showing an improving trend, the distance between dyslexic and typical readers remains greater.

One view is that children are phonologically accurate but that their phonological processing is slow. Mayringer & Wimmer (2000) reported that Austrian dyslexic children are consistently deficient in a pseudowords learning task.

Figure 10 –- Text reading speed for dyslexics: comparison between mild and moderate-severe dyslexics

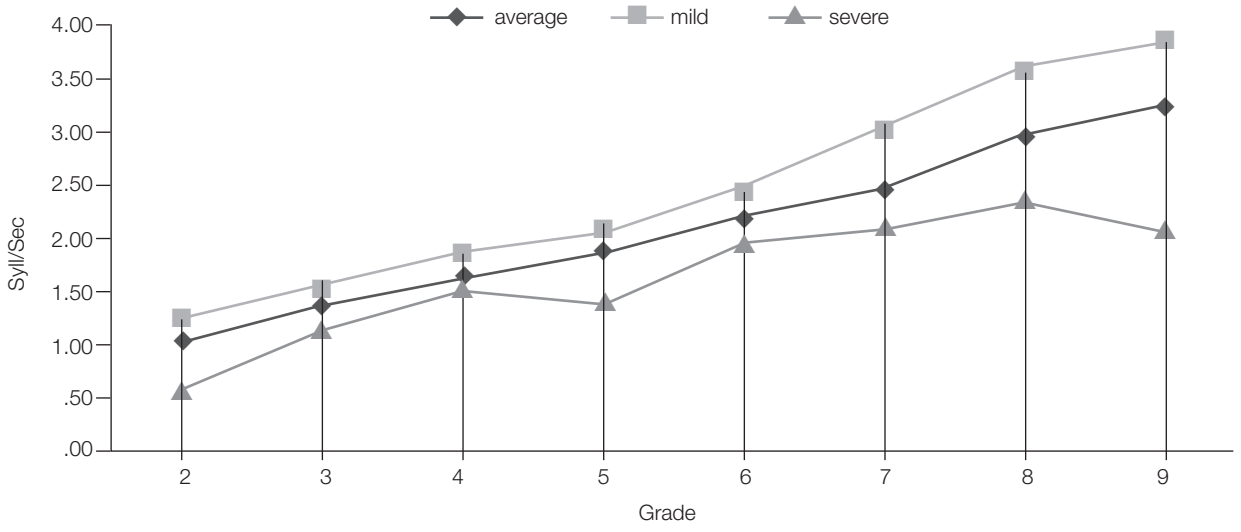
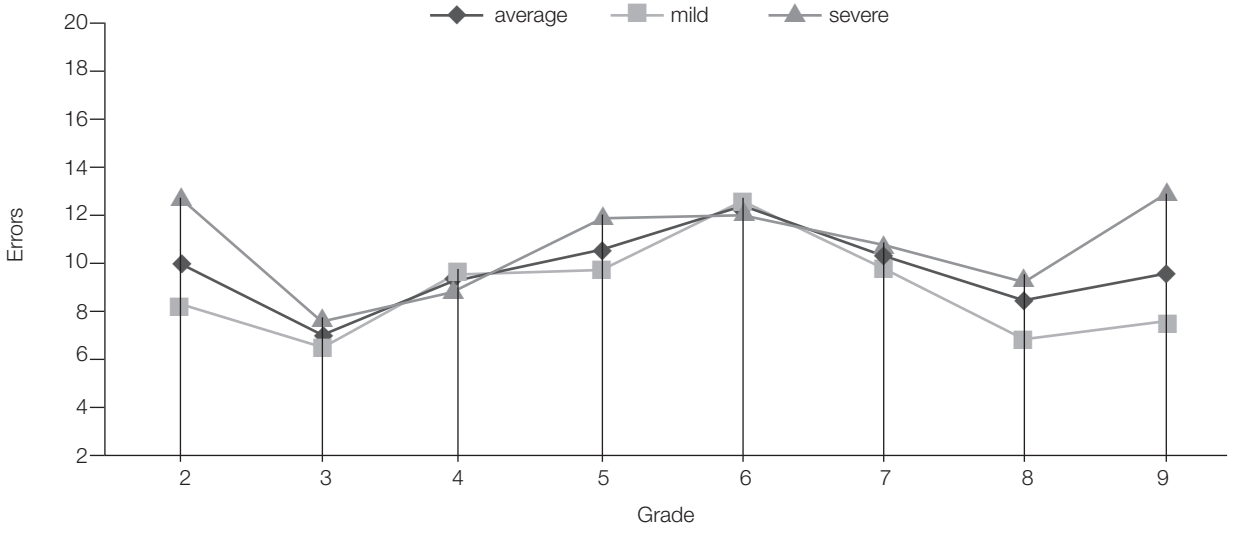


Figure 11 – Accuracy in reading text for dyslexics: comparison between mild and moderate-severe dyslexics



Furthermore, pseudowords reading speed was more impaired in dyslexics than in age- or reading-matched controls than word reading speed. In other words, this view assumes that the reading defect is based on "phonological inefficiency" (Di Filippo, De Luca, Judica, Spinelli & Zoccolotti, 2006).

In relation to the severity of the reading disorder, the data of this study allow to make further considerations on the development of dyslexia. As regards the speed parameter, the data showed that there is a statistically significant difference between the annual average increase of dyslexic mild and medium - severe both in the reading of the words and of the text. In the reading of pseudowords instead there is no statistically significant difference: the phonological decoding seems to be more compromised, regardless of the characteristics of the language system (Rack, Snowling & Olson, 1992; Vellutino, Fletcher, Snowling & Scanlon, 2004; Ziegler et al., 2003).

It is known that what allows a fast and fluid reading is the use of their lexical knowledge and this applies to both languages in regular spelling, such as Italian, which for the opaque like English. The lexicality effect, the frequency effect, the effect of imaginability and the effect of age of acquisition

Table 1 – Estimated parameters for the linear interpolating functions

	Words reading speed	Pseudo words reading speed	Text reading speed	Words reading fluency	Pseudo words reading fluency	Accuracy in reading words	Accuracy in reading pseudo words	Accuracy in reading text
	AVERAGE REGRESSION LINE FOR DISLEXICS							
Slope	.294	.125	.312	–37.498	–17.045	–1.304	–.964	.120
R^2	.994	.963	.983	.840	.747	.952	.901	.037
***p*-value (test F)**	.000	.000	.000	.001	.006	.000	.000	.647
	REGRESSION LINE FOR MILD DISLEXICS							
Slope	.369	.141	.391	–31.466	–12.736	–1.314	–1.088	.019
R^2	.994	.961	.980	.878	.853	.863	.837	.001
***p*-value (test F)**	.000	.000	.000	.001	.001	.001	.001	.953
	REGRESSION LINE FOR MODERATE-SEVERE DYSLEXICS							
Slope	.217	.115	.229	–51.960	–26.599	–1.376	–.863	.182
R^2	.883	.840	.869	.767	.628	.850	.736	.061
***p*-value (test F)**	.000	.000	.000	.001	.006	.000	.000	.000
	DIFFERENCES BETWEEN THE SLOPE FOR MILD DISLEXICS AND THE SLOPE FOR MODERATE-SEVERE DYSLEXICS							
	.153	.025	.161	20.494	13.863	.062	–.226	–.162

are documented since the early years of schooling (Tressoldi, 1996; Zoccolotti & Burani, 2010). For example, the effect of lexicality and stimulus length was studied by Di Filippo et al. (2006) in 32 third- and fourth-grade Italian dyslexics and in 86 age-matched controls and the results were analyzed in terms of raw reaction time (RT). The results showed that in terms of RT, dyslexics exhibited a larger difference between words and pseudowords (lexicality effect) and between short and long stimuli (length effect) than typical readers. This pattern indicates that stimulus length has a specific role in Italian dyslexics' reading deficit. Ziegler et al. (2003) investigated reading characteristics of dyslexic children in regular and less regular orthographies and he considered three critical marker effects of the reading process such as effects of lexicality, length and large orthographic units. The results of this study clearly showed that the similarities between orthographies were far bigger than their differences: English and German dyslexics exhibited a reading speed deficit, a nonword reading deficit and an extremely slow and serial phonological decoding mechanism. These problems were of similar size across orthographies and persisted. The bottleneck of the dyslexic children in both countries seems to lie in the establishment of basic phonological recoding procedures. (Ziegler et al., 2003).

In addition, the results of this study have shown that the average annual increase in decoding speed in the reading of pseudowords of mild and medium - severe dyslexic is not statistically significant. This data confirms the available literature (Rack et al., 1992; Van den Broeck & Geudens, 2012; Ziegler et al., 2003). Infact, the size of the phonological decoding deficit can be estimated by comparing the difference between word and pseudo-word reading across different groups of readers. The words were read faster and more accurately than pseudowords (Rack et al., 1992; Ziegler et al., 2003). The deficit of pseudowords then it would seem not only characterize as dyslexia regardless of the language system, but it would seem the core deficits even in milder forms of dyslexia. In addition, some studies highlight how even the dyslexic adults compensated continue to experience difficulties in this task (Ghidoni, 2011; Hatcher et al., 2002; Martino et al., 2011).

As for the accuracy parameter, the dyslexic group mild improves constantly both in the reading of the words and of pseudowords, thereby reducing their distance from the average, while in the reading of the text show a trend that is not linear. This trend may depend on the increasing length and greater linguistic complexity of the tracks to read. It can therefore be assumed that increasing the reading speed will also increase the number of errors committed.

The group of medium - severe dyslexic instead shows a non-linear trend in all proposed stimuli. Based on these data it is possible to assume that the severity of the reading disorder affects the correctness greater extent than in the mild dyslexic group.

LIMITATION AND FUTURE RESEARCH

An important limitation of this study concerns the distribution of the sample that does not cover all the classes of the secondary school, but only the first class (level 9). So, this distribution does not allowed to make a comparison on the development of the reading ability throughout compulsory education. Clinically the results of this study permit some observations.

First of all, the slowness in decoding is a critical marker of the reading disorder.

Moreover, the absence of statistically significant difference between words reading and text reading supposes that lexical strategy does not sufficiently support reading decoding but also some aspects of text comprehension. Considering that there are reported more comprehension difficulties in the upper secondary school respect to the primary school, it would be necessary to analyze if there is a relationship between the severity of the reading disorder and the text comprehension disorder. This is certainly a future aim of research.

CONCLUSIONS

The aim of this study was to investigate the evolution of reading disorder in the course of compulsory schooling and see if there are different evolutionary trajectories in relation to the severity of the reading disorder. The available literature on the subject considers dyslexia a persistent disorder over the years of compulsory education. In regular orthographies, the critical aspect is the reading speed (speed dyslexia; Wimmer, 1993), whereas decoding accuracy increases. In other words, children with dyslexia improve their overall reading ability, but they are still quite distant from their normal-reading peers

Data collected in the present work confirm that the gap between dyslexics and normal readers persists and that the decoding deficit concerning speed and accuracy develops differently. Reading remains a hard task for dyslexics since they show a slower and less fluent reading than typical readers.

There are differences also in the development of the reading profile between mild and severe dyslexics. The performance trajectory for the moderate-severe dyslexics shows some plateaus and a decrease in performances in the last year analyzed (1st upper secondary school) while the trajectory for the mild dyslexics always show increases in performances. All subjects show a steady increase in word and text reading speed and a slower improvement in pseudo-word decoding.

In terms of accuracy, the trajectory is less smooth. The mild dyslexics group outperforms the moderate-severe dyslexics only in some school years. In other years, the performances are similar.

These findings are consistent with those of other studies on the subject (Holopainen et al., 2001; Jimenez, 2012; Lyytinen et al., 2006; Shaywitz et al., 1999; Stella et al., 2010; Tressoldi et al., 2001; Tucci et al., 2013; Van der Leij et al., 2013), confirming that the critical sign of the disorder remains the reading speed.

References

ARINA, S., IERVOLINO, I. & STELLA, G. (2013). Prima raccolta di dati normativi per la valutazione della dislessia evolutiva negli adolescenti su un campione di scuola secondaria di secondo grado. *Dislessia*, *10*(1), 9-38.

BRIZZOLARA, D., CHILOSI, A.M. & GASPERINI, F. (2011). La dislessia evolutiva: Modelli fisiopatologici e modalità di diagnosi e trattamento. *Prospettive in Pediatria*, *41*(163), 143-149.

BURANI, C., MARCOLINI, S. & STELLA, G. (2002). How early does morpholexical reading develop in readers of a shallow orthography? *Brain and Language*, *81*, 568–586.

CAMPANINI, S., BATTAFARANO, R. & IOZZINO, R. (2010). Evoluzione naturale della lettura del brano, delle liste di parole e non parole e della comprensione del testo in dislessici mai trattati. *Dislessia*, *7*(2), 165-179.

CARAVOLAS, M., LERVAG, A., DEFIOR, S., MALKOVA, S.M. & HULME, C. (2013). Different patterns but equivalent predictors of growth in reading in consistent and inconsistent orthographies. *Psychological Science*, *24*(8), 1398-1407.

CIUFFO, M., GAGLIANO, A., INGRASSIA, M., GERMANÒ, E., BENEDETTO, L., ANGELINI, D., GHIDONI, E. & STELLA, G. (2014). Uno studio pilota per la rilevazione della velocità di lettura in modalità silente: Implicazioni per la valutazione di adolescenti e adulti con dislessia. *Dislessia*, *11*(3), 341-354.

COLTHEART, M., RASTLE, K., PERRY, C., LANGDON, R. & ZIEGLER, J.C. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204-256.

CONSENSUS CONFERENCE (2007). *Disturbi evolutivi specifici di apprendimento: Raccomandazioni per la pratica clinica definite con il metodo della Consensus Conference*. Milano, 2007.

CONSENSUS CONFERENCE (2011). *Disturbi specifici dell'apprendimento*. Roma 6-7 dicembre 2010. Istituto Superiore di Sanità, http://www.snlg-iss.it

CORNOLDI, C. & COLPO, G. (1995). *Nuove prove di lettura MT per la scuola elementare e media*. Firenze: Giunti O.S. Organizzazioni Speciali.

CORNOLDI, C. & COLPO, G. (2012). *Nuove prove di lettura MT per la scuola elementare e media*. Firenze: Giunti O.S. Organizzazioni Speciali.

CORNOLDI, C., PRABALDI, A., FRISO, G., GIACOMIN, A., GIOFRÈ, D. & ZACCARIA S. (2010). *Prove MT avanzate-2 di lettura e matematica per il biennio della scuola superiore di II grado*. Firenze: Giunti O.S. Organizzazioni Speciali.

CORNOLDI, C. & TRESSOLDI, P.E. (2014). Linee guida per la diagnosi dei profili di dislessia e disortografia previsti dalla legge 170: Invito a un dibattito. *Psicologia clinica dello sviluppo*, *XVIII*(1), 75-142.

DELLAI, F., LIPPARINI, S., CORNOLDI, C. & ENGLARO, G. (2014). Il corso longitudinale della dislessia: risultati di un follow-up a 7 anni di distanza su un gruppo di studenti con dislessia. *Dislessia*, *11*(3), 281-294.

DI FILIPPO, G., DE LUCA, M., JUDICA, A., SPINELLI, D. & ZOCCOLOTTI, P. (2006). Lexicality and stimulus lenght effects in Italian dyslexics: role of the overaddivity effect. *Child Neuropsychology*, *12*(2), 141-149.

GHIDONI, E. (2011). Dislessia negli adulti: Una rassegna delle conoscenze. In E. Genovese, E. Ghidoni, G. Guaraldi e G. Stella (a cura di), *Dislessia nei giovani adulti: Strumenti compensativi e strategie per il successo*. Trento: Erickson.

GOSWAMI, U. & ZIEGLER, J.C. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3-29.

GOSWAMI, U. & ZIEGLER, J.C. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science*, *9* (5), 429-436.

HATCHER, J., SNOWLING, M.J. & GRIFFITHS, Y.M. (2002). Cognitive assessment of dyslexic students in higher education. *British Journal of Educational Psychology*, *72*, 119–133.

HOLOPAINEN, L., AHONEN, T. & LYYTINEN, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities*, *34*, 401-413.

JIMENEZ, J.E. (2012). *Dyslexia en espanol*. Madrid: Piramide.

JIMENEZ, J.E. & HERNANDEZ-VALLE, I. (2000). Word identification and reading disorders in the Spanish language. *Journal of Learning Disabilities*, *33*, 44-60.

KASPRZYK, K.D., DUNKAN, G.J., KALTON, G. & SING, M.P. (1989). New York: John Wiley and Sons.

LAMI, L., PALMIERI, A., SOLIMANDO, M.C. & PIZZOLI, C. (2008). Evoluzione del profilo di lettura nella dislessia: Studio longitudinale su un gruppo di dislessici divenuti giovani adulti. *Dislessia*, *5*(1), 7-17.

LEINONEN, S., MULLER, K., LEPPANEN, P.H.T., ARO, M., AHONEN, T. & LYYTINEN, H. (2001). Heterogenity in adult dyslexic readers: Relating processing skills to the speed and accuracy of oral text reading. *Reading and writing: An interdisciplinary journal*, *14*, 265-296.

LEONARD, L.B. (2009). Some reflections on the study of children with specific language impairment. *Child language teaching and therapy*, *25*, 2, 169-171.

LORUSSO, M.L., TORALDO, A. & CATTANEO, C. (2006). Parametri "tempo" e "velocità" per la misurazione della rapidità di lettura. *Dislessia*, *3*(3), 263-282.

LOSITO, N., TRESSOLDI, P.E. & CORNOLDI, C. (2014). Punti z o percentili? Sillabe/secondo, tempo complessivo tempo/sillaba? Come valutare la rapidità nelle prove di lettura. *Dislessia*, *11*(3), 295-311.

LYYTINEN, H., ERSKINE, J., TOLVANEN, A., TORPPA, M., POIKKEUS, A.M. & LYYTINEN, P. (2006). Trajectories of reading development: A follow-up from birth to school age of children with and without risk for dyslexia. *Merrill-Palmer Quarterly*, *52*(3), 514-546.

MARTELLI, M., DI FILIPPO, G., SPINELLI, D. & ZOCCOLOTTI, P. (2009). Crowding, reading and developmental dyslexia. *Journal of Vision*, *9*, 14.

MARTINO, M.G., PAPPALARDO, F., RE, A.M., TRESSOLDI, P.E., LUCANGELI, D. & CORNOLDI, C. (2011). La valutazione della dislessia nell'adulto. Un contributo alla standardizzazione della Batteria dell'Università di Padova. *Dislessia*, *8*(2), 119-134.

MAUGHAN, B., MESSER COLLISHAW, J.S., PICKLES, A., SNOWLING, M., YULE, W. & RUTTER, M. (2009). Persistence of literacy problems: Spelling in adolescence and at mid-life. *Journal of Child Psychology and Psychiatry*, *50*(8), 893–901.

MAYRINGER, H., & WIMMER, H. (2000). Pseudoname learning by German-speaking children with dyslexia: Evidence for a phonological learning deficit. *Journal of Experimental Child Psychology*, *75*, 116–33.

McCANDLISS, B.D. (2009). Helping dyslexic children attend to letters within visual word forms. *PNAS*, *109*, 11064-11065.

McLAUGHLIN, M.J., SPEIRS, K.E. & SHENASSA, E.D. (2014). Reading disability and adult attained education and income: evidence from a 30-year longitudinal study of a population-based sample. *Journal of Learning Disabilities*, *47*(4), 374-386.

MORLINI, I., STELLA, G. & SCORZA, M. (2013). A new procedure to measure children's reading speed and accuracy in Italian. *Dyslexia*, *20*, 54-73.

MORLINI, I., STELLA, G. & SCORZA, M. (2014). Some experimental results on the role of speed and accuracy of reading in psychometric tests. In S. Mignani & F. Crescenzi, *Statistical Methods and Applications from a Historical Perspective.* Berlin: Springer DEU.

MORLINI, I., STELLA, G. & SCORZA, M. (2015). Assessing decoding ability: The role of speed and accuracy and a new composite indicator to measure decoding skill in elementary grades. *Journal of Learning Disabilities*, *48*, 176-195.

NATION, K., COCKSEY, J., TAYLOR, J.S.H. & BISHOP, D.V.M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry*, *51*(9), 1031-1039.

PANEL DI AGGIORNAMENTO E REVISIONE DELLA CONSENSUS CONFERENCE DSA (2011). *Raccomandazioni cliniche sui DSA: Risposte a quesiti*. DSA, Documento d'intesa, PARCC 2011, http://www.lineeguidadsa.it

PAULESU, E., DÉMONET, J.F., FAZIO, F., McCRORY, E., CHANOINE, V., BRUNSWICK, N., CAPPA, S.F., COSSU, G., HABIB, M., FRITH, C.D. & FRITH, U. (2001). Dyslexia: Cultural diversity and biological unity. *Science*, *291*, 2165 – 2167.

PETERSON, R.L. & PENNINGTON, B.F. (2012). Developmental dyslexia. *The Lancet*, *379*(9830), 1997-2007.

RACK, J.P., SNOWLING, M.J. & OLSON, R.K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly*, *27*, 29-53.

RAMUS, F. (2006). A neurological model of dyslexia and other domain-specific developmental disorders with an associated sensorimotor syndrome. In G.D. Rosen (Ed.), *The Dyslexic Brain: New Pathways in Neuroscience Discovery*. Mahwah, NJ: Lawrence Erlbaum Associates.

RAMUS, F., ROSEN, S., DAKIN, S.C., DAY, B.L., CASTELLOTE, J.M., WHITE, S. & FRITH, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, *126*, 841-865.

RE, A.M., LUCANGELI, D. & CORNOLDI, C. (2011). Nuove evidenze sulle difficoltà di lettoscrittura di studenti universitari con una storia di dislessia. In E. Genovese, E. Ghidoni, G. Guaraldi e G. Stella (a cura di), *Dislessia nei giovani adulti: Strumenti compensativi e strategie per il successo*. Trento: Erickson.

RE, A.M., TRESSOLDI, P.E., CORNOLDI, C. & LUCANGELI, D. (2011). Which tasks best discriminate between dyslexic university student and controls in a trasparent language? *Dyslexia*, *17*, 227-241.

SARTORI, G., JOB, R. & TRESSOLDI, P.E. (1995). Batteria per la valutazione della dislessia e disortografia evolutiva. Firenze: Giunti O.S. Organizzazioni Speciali.

SARTORI, G., JOB, R. & TRESSOLDI, P.E. (2007). Batteria per la Valutazione della Dislessia e della Disortografia Evolutiva-2/ DDE-2. Nuova edizione con norme aggiornate. Firenze: Giunti O.S. Organizzazioni Speciali.

SCORTICHINI, F., GASPERINI, C., SCORZA, M., BONI, C.D. & STELLA, G. (2015). The European Battery for reading assessment: Interlingual comparative study of reading abilities. *BPA Applied Psychology Bulletin*, *274*, 49-63.

SCORZA, M., BONI, C.D., ZANZURINO, G.G.F., SCORTICHINI, F., MORLINI, I. & STELLA, G. (2015). Decoding Automaticity in Reading Process and Practice. How much influence does summer vacation have on children's reading abilities in primary school? *BPA Applied Psychology Bulletin*, *273*, 17-29.

SERRANO, F. & DEFIOR, S. (2008). Dyslexia speed problems in a transparent orthography. *The International Dyslexia Association*, *58*, 81-95.

SHAYWITZ, S.E., FLETCHER, J.M., HOLAHAN, J.M., SHNEIDER, A.E., MARCHIONE, K.E., STUEBING, K.K., FRANCIS, D.J., PUGH, K.R. & SHAYWITZ, B.A. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study at adolescence. *Pediatrics*, *104*, 6.

STELLA, G. (2013). Disturbi specifici del linguaggio e dell'apprendimento. In S. D'Amico & A. Devescovi (a cura di), *Psicologia dello sviluppo del linguaggio*. Bologna: Il Mulino.

STELLA, G. & CERRUTI BIONDINO, E. (2002). La dislessia evolutiva lungo l'arco della scolarità obbligatoria. In S. Vicari & M.C. Caselli (a cura di), *I disturbi dello sviluppo, neuropsicologia clinica e ipotesi riabilitative*. Bologna: Il Mulino

STELLA, G., DI BLASI, F., GIORGETTI, W. & SAVELLI, E. (2003). *La valutazione della dislessia*. Troina (EN): Città Aperta Edizioni.

STELLA, G., SAVELLI, E., SCORZA, M. & MORLINI, I. (2010). La dislessia evolutiva lungo l'arco della scolarità obbligatoria. In S. Vicari & M.C. Caselli (a cura di), *Neuropsicologia dello sviluppo*. Bologna: Il Mulino.

STELLA, G. & TINTONI, C. (2007). Indagine e rilevazione sulle abilità di lettura nelle scuole secondaria di secondo grado. *Dislessia*, *4*(3), 271-285.

TRESSOLDI, P.E. (1996). Evoluzione della lettura e della scrittura dalla seconda elementare alla terza media. *Età evolutiva*, *53*, 43-55.

TRESSOLDI, P.E., LORUSSO, M.L., BRENBATI, F. & DONINI, R. (2007). Fluency remediation in dyslexic children: Does age make a difference? *Dyslexia*, *14*, 142-152.

TRESSOLDI, P.E., STELLA, G. & FAGGELLA, M. (2001). The development of reading speed in Italians with dyslexia. *Journal of Learning Disabilities*, *34*(5), 414-417.

TRESSOLDI, P.E. & VIO, C. (2007). Sillabe al secondo o secondi per sillaba: Qual è il problema? *Dislessia*, *4*(1), 7-11.

TRESSOLDI, P.E., VIO, C. & IOZZINO, R. (2007). Efficacy of an intervention to improve fluency in children with developmental dyslexia in a regular orthography. *Journal of Learning Disabilities*, *40*(3), 203-209.

TUCCI, R., SAVOIA, V., MERELLA, A. & TRESSOLDI, P.E. (2013). La dislessia evolutiva lungo l'arco della scolarità obbligatoria – Una replica dello studio di Stella et al. del 2001. *Dislessia*, *10* (3), 281-294.

TUCCI, R. & TRESSOLDI, P.E. (2009). Lo sviluppo della lettura e la definizione di dislessia in età adolescenziale e adulta – Una proposta. *Dislessia*, *6* (2), 269-279.

UNDHEIM, A.M. (2009). A thirteen-year follow-up study of young Norwegian adults with dyslexia in childhood: reading development and educational levels. *Dyslexia*, *15* (4), 291-303.

VAN DEN BROECK, W. & GEUDENS, A. (2012). Old and new ways to study characteristics of reading disability: The case of the nonword-reading deficit. *Cognitive Psychology*, *65*, 414-456.

VAN DER LEIJ, A. (2013). Dyslexia and early intervention: What did we learn from the Dutch Dyslexia Programme? *Dyslexia*, *19* (4), 241-255.

VAN DER LEIJ, A., VAN BERGEN, E., VAN ZUIJEN, T., DE JONG, P., MAURITS, N. & MAASSEN, B. (2013). Precursors of developmental dyslexia: An overview of the longitudinal Dutch Dyslexia Programme study. *Dyslexia*, *19*, 191-213.

VAN DER LEIJ, A. & VAN DAAL, V.H.P. (1999). Automatization aspects of dyslexia: Speed limitations in word identification, sensitivity to increasing task demands and orthographic comprehension. *Journal of Learning Disabilities*, *32*, 417-428.

VELLUTINO, F.R., FLETCHER, J.M., SNOWLING, M.J. & SCANLON, D.M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatric*, *45* (1), 2-40.

WIMMER, H. (1993). Characteristic of developmental dyslexia in a regular writing system. *Applied psycholinguistic*, *14*, 1-33.

WIMMER, H. & MAYRINGER, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology*, *94*, 272–277.

ZIEGLER, J.C., PERRY, C., MA-WYATT, A., LADNER, D. & SCHULTE-KORNE, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology*, *86*, 169-193.

ZOCCOLOTTI, P. & BURANI, C. (2010). Apprendimento e disturbi di lettura in ortografie trasparenti. In S. Vicari & M.C. Caselli (a cura di), *Neuropsicologia dello sviluppo*. Bologna: Il Mulino.

ZOCCOLOTTI, P., DE LUCA, M., DI FILIPPO, G., JUDICA, A. & MARTELLI, M. (2008). Reading development in an orthographically regular language: Effects of length, frequency, lexicality and global processing ability. *Reading and Writing*, *22*, 1053-1079.

ZORZI, M., BARBIERO, C., FACOETTI, A., LONCIARI, I., CARROZZI, M., MONTICO, M., BRAVAR, L., GEORGE, F., PECH-GEORGEL, C. & ZIEGLER, J.C. (2012). Extra-large letter spacing improves reading in dyslexia. *PNAS*, *109*, 11455–11459.

The Italian version of the Job Crafting Scale (JCS)

Roberto Cenciotti[1], Laura Borgogni[1], Antonino Callea[2], Lara Colombo[3], Claudio Giovanni Cortese[3], Emanuela Ingusci[4], Mariella Miraglia[5], Margherita Zito[3]

[1] Department of Psycology, Sapienza University of Rome
[2] Department of Human Science, LUMSA University of Rome
[3] Department of Psycology, University of Turin
[4] Department of History, Society and Human Studies, University of Salento, Lecce
[5] Norwich Business School, University of East Anglia, Norwich

ABSTRACT. Il job crafting si riferisce alle azioni messe in atto dagli individui al fine di adattare le richieste e le risorse lavorative alle proprie preferenze. Questo contributo presenta la versione italiana della Job Crafting Scale, un questionario per la misurazione dei comportamenti di job crafting, che comprende tre fattori: aumentare le risorse strutturali, aumentare le risorse sociali e aumentare le richieste sfidanti. Le caratteristiche psicometriche del questionario sono in linea con quelle descritte in letteratura, e la propensione a mettere in atto comportamenti di job crafting risulta correlata all'autoefficacia, al work engagement e alla prestazione.

SUMMARY. Job crafting refers to actions carried out by workers in order to bring their job demands and job resources at a preferred level. Crafting behaviors are measured by the Dutch Job Crafting Scale (JCS). The Italian version of the JCS includes the following three positive factors: increasing structural job resources, social job resources and challenging job demands. To assess the factorial validity of the scale, an exploratory factor analysis (N=311) and confirmatory factor analyses (N=410) were performed. Convergent and criterion validity were investigated through correlations with other variables. Factor analyses showed a good three-factor structure, in line with the literature. Moreover, as expected, job crafting behaviors were correlated with work self-efficacy, work engagement and job performance. Results suggest that the Italian version of the JCS can be reliably used to measure job crafting.

Keywords: *Iob crafting, Job demands-resources model, Scale adaptation*

INTRODUCTION

During the last decades, the economic global crisis has modified the labour market and forced companies to improve their abilities and know-how to be more competitive. These constant and rapid changes have directly and indirectly involved workers and organizations (Callea, Urbini, Ingusci, & Chirumbolo, 2014), and required them greater flexibility and stronger personal initiative. In this scenario, it has become more urgent to develop and improve new strategies to facilitate individuals' successful coping with the turbulent context. These strategies can be implemented by managers, through interventions aimed at adapting the organization to external modifications (Petrou, 2013; Petrou, Demerouti, & Häfner, 2015), but also generated by the employees themselves. Indeed, research (e.g. Van den Heuvel, Demerouti, Bakker & Schaufeli, 2010) has shown that workers are able to assume a proactive role in remolding their work activities and crafting their job, to activate the desired changes. Therefore, job crafting results of critical importance because it can represent an individual strategy to promote the best conditions for the future. According to Tims, Bakker and Derks (2012), job crafting involves self-initiated changes and behaviors that employees perform in order to adjust their jobs with their preferences, motivations and needs.

The authors (Tims et al., 2012) inscribe the concept within the job demands-resources (JD-R) theoretical framework, which considers two broad classes of processes at work (job demands and job resources) in the development of well-being and performance (Bakker & Demerouti, 2014). On the one hand, job demands are those aspects of the job that require a physical and psychological (cognitive or emotional) effort. Examples of job demands are heavy workload, emotionally demanding interactions with others, or high responsibility. Considering their effects on workers' job outcomes, job demands can be distinguished between challenging demands (i.e., obstacles that workers have to overcome to learn and achieve goals) and hindering demands (i.e., needless requests that impede worker's personal growth and goal achievement) (Bakker & Sanz-Vergel, 2013). On the other hand, job resources are those aspects of the job that are functional to achieve work goals, reduce the physiological and psychological cost associated to job demands, and increase skills learning and development. Examples of job resources are job autonomy or performance feedback.

In this perspective, job crafting is defined as the changes that employees may make to balance their job demands and job resources with their personal abilities and needs (Tims & Bakker, 2010; Tims et al., 2012). Within this conceptualization, Tims et al. (2012) proposed three broaden dimensions of job crafting: increasing job resources, increasing challenging job demands and decreasing hindering job demands. Increasing job resources can result in both positive organizational and individual outcomes, such as work engagement and job satisfaction (Zito, Cortese & Colombo, 2015). Furthermore, optimizing job resources may enhance individual well-being, because they allow employees to protect themselves from exhaustion, sustain their existing resources, and achieve expected outcomes, in line with the Conservation of Resources (COR) theory (Petrou et al., 2015). Increasing challenging job demands can enable individuals to pursue more difficult goals, improve their skills, and avoid boring jobs or repetitive tasks that can reduce the energy and effort at work. Finally, decreasing hindering job demands depicts those employees' behaviors aimed at reducing the emotionally, mentally and physically demanding aspects of the job (e.g., relational stressors) that can limit them in achieving their performance (Petrou, Demerouti, Peeters, Schaufeli & Hetland, 2012). All in all, crafting behaviors represent a very promising strategy to foster employee-organization fit as well as organizational effectiveness.

From an empirical standpoint, job crafting is measured by the Dutch Job Crafting Scale (JCS) developed by Tims et al. (2012). To test the psychometric characteristics of the JCS, the authors conducted three separate studies in the Netherlands (total sample N=1,181). In study 1, they performed an explorative test on the initial 42 items of the JCS and found a four-factor structure instead of the proposed three-factor one, after deleting 21 items with low or ambiguous factor loadings. Study 2 confirmed this four-factor structure on the remaining 21 items. The broader dimension "increasing job resources" was split in two sub-dimensions: (a) structural job resources, referred to organizational resources (e.g., opportunities for development, autonomy and variety), and (b) social job resources, referred to support from colleagues or supervisors (e.g., social support, feedback and coaching). With regard to convergent validity, increasing structural job resources, increasing social job resources and increasing challenging job demands correlated positively with proactive personality and personal initiative (considered as active constructs) and negatively with cynicism (considered, indeed, as an inactive

construct), while decreasing hindering job demands showed a positive and significant correlation only with cynicism. Study 3 examined the criterion validity of the scale and reported that increasing structural job resources, increasing social job resources and increasing challenging job demands were positively correlated with work engagement, employability and performance, while decreasing hindering job demands was not significantly associated with any of these variables.

In this sense, the study (Tims et al., 2012) revealed an evident difference between the first three factors of JCS on the one hand, which are oriented toward a positive direction of *increasing* (job resources and challenging demands), and the fourth factor on the other hand, which is oriented toward a direction of *reducing* (hindering demands). The latter dimension, in fact, showed a peculiar pattern of correlations with outcome variables, different from that of the *increasing* dimensions. In a study by Bakker, Tims and Derks (2012), which examined the role of proactive personality in predicting work engagement and job performance, job crafting was operationalized through the three increasing factors (thus, excluding the behaviors related to decreasing hindering job demands), resulting in a variable that mediates the relationship between proactive personality and work engagement. A further recent study (Tims, Bakker & Derks, 2015) confirmed the difference between the increasing and decreasing dimensions. Indeed, only decreasing hindering job demands did not correlate with work engagement and OCB nor it lead to motivation.

Aims

Based on the aforementioned literature, the present study aims to provide a first psychometric evaluation of the Italian version of the JCS, including the three job crafting dimensions oriented in the positive direction of "increasing": increasing structural job resources, increasing social job resources, and increasing challenging job demands. This general purpose will be declined in three specific aims: (1) to test the factorial validity and reliability of the Italian JCS; (2) to investigate its convergent validity, by analyzing the relation between job crafting and work self-efficacy, representing individual proactivity (see Tims et al., 2012); (3) to analyze the criterion validity, by exploring the relations of crafting behaviors with work engagement and job performance, in line with Tims et al. (2012).

A contribution to the validation of the Italian JCS seems necessary in light of the lack, to our knowledge, of an Italian job crafting measure. Therefore, the present study can fill the gap and promote in Italy more empirical research on the dynamics and consequences of job crafting.

METHODS

Participants

To perform an exploratory factor analysis and a confirmatory factor analysis, two heterogeneous samples of Italian workers were used for the research. The first sample was composed by 311 participants from several organizations: 51.4% female, average age 40 years (SD = 11.4), average seniority 12 years (SD = 10.4). The second sample included 410 participants from a large service organization: 51.1% male, average age 44 years (SD = 9.6), average seniority 14 years (SD = 14.3).

Measures

- *Job crafting*. We used the Italian version of the JCS, consisting of the three increasing dimensions (Bakker et al., 2012): increasing structural job resources (five items, e. g. "I try to develop my capabilities"), increasing social job resources (five items, e. g. "I ask my supervisor to coach me") and increasing challenging job demands (five items, e. g. "When an interesting project comes along, I offer myself proactively as project co-worker"), for a total of 15 items. All items were translated (from English to Italian) and back-translated (from Italian to English) with the help of an English mother tongue speaker. The result was a good correspondence between items. The investigation of validity and reliability of the scale is an aim of the present study. Items were measured on a 7-point frequency scale, ranging from 1 = *Never* to 7 = *Always*.
- *Work self-efficacy*. We used a monofactorial work self-efficacy scale created and validated in the Italian context (Borgogni, Dello Russo, Petitta & Vecchione, 2010). The scale consists of seven statements assessing the beliefs of being able to handle job responsibilities, challenging situations and coordination with colleagues (e.g. "In my

work I am confident I can generate new ideas in order to deal with organizational demands", α = .92). The statements were measured on a 7-point scale, ranging from 1 = *Cannot do at all* to 7 = *Highly certain can do*.

– *Work engagement*. We used the validated Italian version of the UWES-9 (Balducci, Fraccaroli & Schaufeli, 2010). The scale entails three factors: vigor, measured by three items (e.g. "At my work, I feel bursting with energy", α = .83); dedication, measured by three items (e.g. "I'm proud of the work that I do", α = .83); and absorption, also measured by three items (e.g. "I am immersed in my job", α = .71). Items were answered using a 7-point frequency scale, ranging from 1 = *Never* to 7 = *Always*.
– *Job performance ratings*. Supervisors rated their employees' performance through the company's established performance appraisal system. This instrument, developed by the HR department of the organization, assesses performance as a general, unidimensional measure. Employees' performance was measured on a 10-point scale (labels: 1 = *Inadequate*; 2-3 = *Improvable*; 4-6 = *Average*; 7-9 = *Elevated*; 10 = *Beyond expectations*).

Procedure

Part of the data on the Italian JCS was collected through a paper-and-pencil questionnaire in the first sample. Afterwards, the second sample filled in an online questionnaire that measured job crafting, work self-efficacy and work engagement.

Participation in the study was voluntary, and a cover letter informed participants about how to complete the paper-and-pencil or online questionnaires (for the first and second samples, respectively) and about data confidentiality. Moreover, for the second sample, supervisory performance ratings were provided at the end of the year by the Human Resource (HR) department of the organization. In order to match the answers provided by each employee with his/her performance ratings, the HR department assigned a code to each participant. The code was used to log in and respond to the online questionnaire. In this way, the HR department knew the name of the employee, his/her code, and the performance rating, but did not know the answers to the questionnaire, whereas the research team knew the code, the answers to the questionnaire, and the performance rating provided by the company, but not the name of employee.

Data analysis

To assess the factorial validity of the Italian JCS (aim 1), first an exploratory factor analysis (EFA) was performed on the first sample through SPSS 20.Principal Axis Factoring extraction method and Promax rotation were used (Kaiser's normalization), since factors were expected to correlate.

Reliability analyses (corrected item-total correlations and Cronbach's alphas) and confirmatory factor analyses (CFA) were performed on the second sample, using Mplus7 (Muthén & Muthén, 2012) for the CFA. To test the model goodness of fit, the following indices were considered: the chi-square value (χ^2); the Comparative Fit Index (CFI); the Tucker-Lewis Index (TLI); the Root Mean Square Error of Approximation (RMSEA); the Standardized Root Mean Square Residual (SRMR). Moreover, to verify the association of job crafting with other relevant variables (aims 2 and 3), its correlations with work self-efficacy, work engagement and job performance were investigated for the second sample, by using Pearson's *r* coefficient.

RESULTS

As regards the EFA, the resulting structure was in line with the scale developed by Tims et al. (2012), with regard to its positive dimensions, and showed three factors (see Table 1): increasing structural job resources (five items), increasing social job resources (five items) and increasing challenging job demands (five items). Factor loadings ranged between |.45| and |.83| for increasing structural job resources, between |.44| and |.87| for increasing social job resources, and between |.57| and |.79| for increasing challenging job demands.

The factor solution absorbs 55% of the total variance. More specifically, increasing structural job resources explained 38% of the variance, increasing social job resources explained 12%, and increasing challenging job demands explained 5%.

Factors reported a correlational pattern quite similar to the one in Tims et al. (2012) study: the higher correlation resulted between increasing structural job resources and increasing challenging job demands (r = .66), followed by the correlations between increasing social job resources and increasing challenging job demands (r = .43), and between increasing structural job resources and increasing social job resources (r = .36).

However, considering the reliability properties of the 15 items (analyzed on the second sample), we found that two

Table 1 – Exploratory factor analysis on the initial 15-item JCS (PAF extraction; Promax rotation; Kaiser's normalization; N = 311)

Item Code	Items	M	SD	Factors		
				STR	SOC	CHA
Str2	Creo le condizioni per crescere professionalmente [I try to develop myself professionally]	5.03	1.38	**.83**	.02	.02
Str1	Creo le condizioni per sviluppare le mie capacità sul lavoro [I try to develop my capabilities]	4.95	1.33	**.82**	.04	.02
Str3	Faccio in modo di imparare nuove cose al lavoro [I try to learn new things at work]	5.49	1.27	**.80**	.01	.04
Str4	Uso a pieno le mie capacità [I make sure that I use my capacities to the fullest]	5.39	1.26	**.69**	–.05	.00
Str5	Decido autonomamente come svolgere il mio lavoro [I decide on my own how I do things]	4.82	1.42	**.45**	–.08	.11
Soc2	Chiedo al mio capo se è soddisfatto del mio lavoro [I ask whether my supervisor is satisfied with my work]	3.45	1.76	–.11	**.87**	–.04
Soc1	Chiedo al mio capo di farmi da "coach" [I ask my supervisor to coach me]	3.80	1.62	.15	**.77**	–.18
Soc4	Chiedo ad altre persone di darmi feedback sulla mia prestazione [I ask others for feedback on my job performance]	3.64	1.73	–.24	**.73**	.21
Soc3	Prendo ispirazione dal mio capo [I look to my supervisor for inspiration]	4.05	1.77	.18	**.66**	–.06
Soc5	Chiedo consigli ai miei colleghi [I ask colleagues for advice]	4.70	1.34	.01	**.44**	.14
Cha4	Mi faccio carico regolarmente di attività "extra", pur non ricevendo alcun compenso per queste [I regularly take on extra tasks even thug I do not receive extra salary for them]	4.57	1.56	–.03	–.08	**.79**
Cha2	Se ci sono delle novità, sono tra i primi ad acquisirle e testarle [If there are new developments, I am one of the first to learn about them and try them out]	4.67	1.45	.04	–.04	**.76**
Cha3	Quando non c'è molto da fare al lavoro, ne approfitto per iniziare nuovi progetti [When there is no much to do at work, I see it as a chance to start new projects]	4.50	1.48	.03	.06	**.66**
Cha1	Quando arriva un progetto interessante, offro proattivamente la mia collaborazione [When an interesting project comes along, I offer myself proactively as project co-worker]	4.99	1.45	.14	.14	**.62**
Cha5	Mi sforzo di rendere il mio lavoro più stimolante riconoscendo tutte le relazioni tra i suoi diversi aspetti [I try to make my work more challenging by examining the underlying relationships between aspects of my job]	4.81	1.35	.26	.00	**.57**

continued on next page

continued

Correlation between factors

	STR	SOC	CHA
STR	1		
SOC	.36	1	
CHA	.66	.43	1

Note. STR = increasing structural job resources; SOC = increasing social job resources; CHA = increasing challenging job demands; M = mean; SD = standard deviation.

items, i.e. Str5 ("I decide on my own how I do things") and Soc5 ("I ask colleagues for advice"), showed a low item-total correlation (.29 and .39, respectively). Accordingly, item Str5 was dropped, since its correlation with the scale (.29) was far below the limiting value (.40). In order to decide whether to maintain item Soc5, a CFA was run on the remaining 14 items. Since the resulting fit indices were not completely adequate[1], item Soc5 was eliminated, obtaining the final 13-item scale (ΔX^2= 61.254, *df* = 12, *p* = .000).

Finally, a CFA was conducted on the posited three-factor model (i. e. Model 1) and its fit compared with several alternative models by testing the change in X^2. These alternative models assumed a two-factor structure, obtained by combining two of the three dimensions (i.e., Models 2, 3 and 4), or a mono-factorial structure (i.e., Model 5, see Table 2). In line with our theoretical assumptions, the three-factor model showed the best fit with the data, suggesting the conformity of the Italian JCS to the scale developed by Tims et al. (2012) and its factorial validity.

All items of the three-factor model (Model 1) loaded only on the hypothesized factors and factor loadings ranged between |.55| and |.82| for increasing structural job resources, between |.53| and |.74| for increasing social job resources, and between |.48| and |.71| for increasing challenging job demands (see Figure 1). Correlations between factors were good. In particular, it has to be noted the elevated correlation between increasing structural job resources and increasing challenging job demands (.92). In this regard, as above mentioned, the fit of a two-factor solution that merged these two dimensions (Model 2) was worse than the fit of the three-factor structure (Model 1, see M2-M1 comparison in Table 2). More specifically, the TLI was lower than the limiting value of .90, making Model 2 not completely acceptable (Tucker & Lewis, 1973). Therefore, despite the high correlation, this study cannot consider the two dimensions of increasing structural job resources and increasing challenging job demands as a unique one.

As regards the reliability statistics (Cronbach's alphas and item-total correlations), they were adequate for each scale, i.e. increasing structural job resources (four items, α = .81, item-total correlations ranging from .52 to .69), increasing social job resources (four items, α = .74, item-total correlations ranging from .43 to .62) and increasing challenging job demands (five items, α = .78, item-total correlations ranging from .45 to .62).

Finally, as expected, the three job crafting dimensions (i.e. increasing structural job resources, increasing social job resources and increasing challenging job demands) were positively correlated with work self-efficacy and work engagement. In particular, correlations among increasing structural job resources and increasing challenging job demands, on the one side, and work self-efficacy and engagement, on the other side, were strong (ranging from .44 and .59), whereas increasing social job resources was more weakly correlated with self-efficacy (*r* = .15) and engagement (*r* = .20). Moreover, increasing challenging job demands and increasing social job resources correlated positively, although modestly (*r* = .19 and .14 respectively), with job performance, whereas increasing structural job resources showed no significant association with performance (see Table 3).

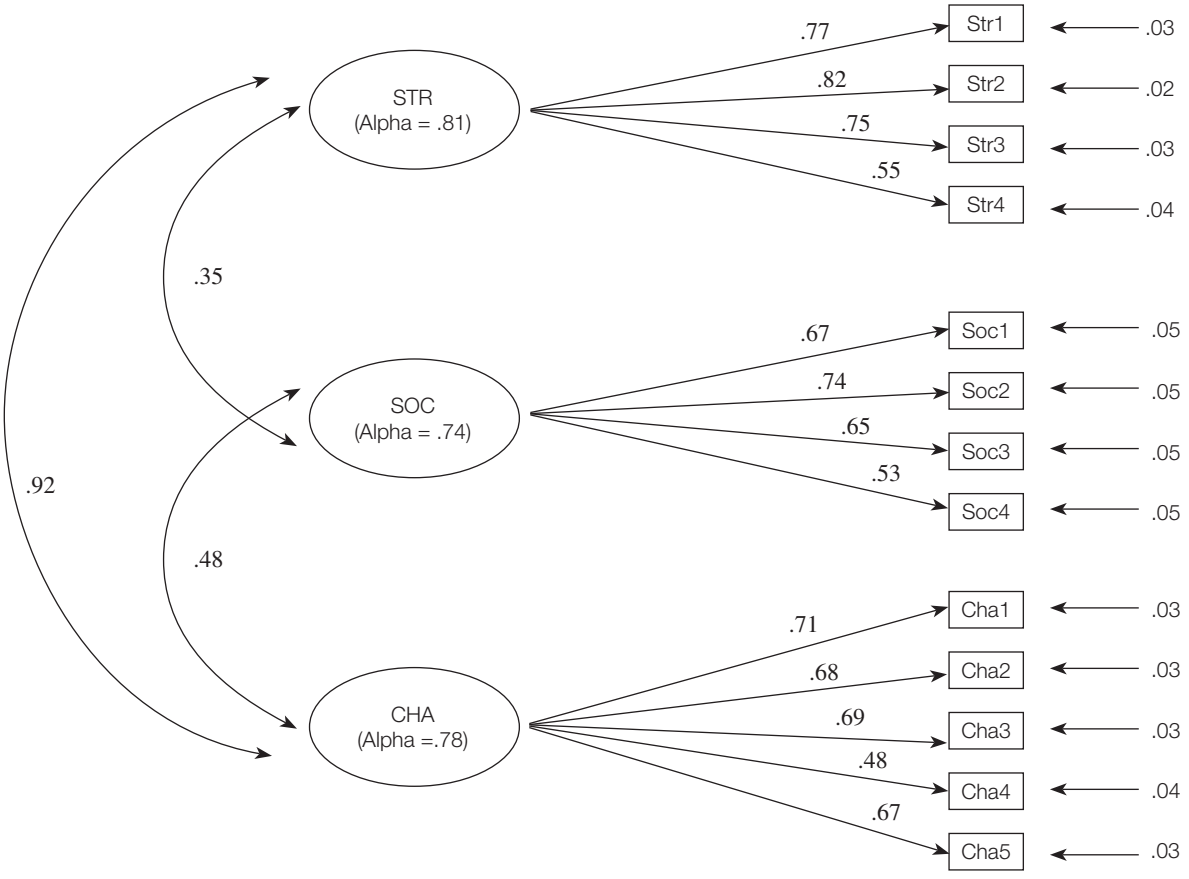
1 X^2 (df = 74) = 268.624, *p* = .000; RMSEA =.08; CFI = .91; TLI = .88; SRMR = .05

Table 2 – Results of the confirmatory factor analysis on the final 13-item JCS: model comparison (N = 410)

MODEL	X^2	df	p	RMSEA	CFI	TLI	SRMR	Model comparison	ΔX^2	df	p
Model 1: 3-Factor Model	207.370	62	.000	.08	.93	.91	.05				
Model 2: 2-Factor Model STR+CHA, SOC	235.097	64	.000	.08	.91	.89	.05	M2-M1	27.727	2	.000
Model 3: 2-Factor Model STR+SOC, CHA	485.545	64	.000	.13	.78	.74	.09	M3-M1	278.175	2	.000
Model 4: 2-Factor Model SOC+CHA, STR	457.523	64	.000	.12	.80	.75	.09	M4-M1	250.153	2	.000
Model 5: 1-Factor Model	494.778	65	.000	.13	.78	.74	.09	M5-M1	287.408	3	.000

Note. STR = increasing structural job resources; SOC = increasing social job resources; CHA = increasing challenging job demands.

Figure 1 – Results of the confirmatory factor analysis and Cronbach's alphas on the final 13-item JCS (N = 410)



Note. STR = increasing structural job resources; SOC = increasing social job resources; CHA = increasing challenging job demands.

Table 3 – Relations to other constructs: Pearson's r coefficients (N = 410)

Dimensions of job crafting	Work self-efficacy	Work engagement	Job performance ratings
STR	.57*	.59*	.10
SOC	.15*	.20*	.14*
CHA	.58*	.44*	.19*

Note. STR = increasing structural job resources; SOC = increasing social job resources; CHA = increasing challenging job demands; * $p<.01$.

CONCLUSIONS

The overall purpose of the study was to provide first psychometric evaluations of the Italian version of the JCS developed by Tims et al. (2012), operationalized by using the three dimensions oriented in the positive direction of *increasing* (i.e., increasing structural job resources, increasing social job resources, and increasing challenging job demands), as suggested by literature (Bakker et al., 2012).

As expected, the exploratory factor analysis revealed a three-factor structure. The content of each factor was in line with our theoretical assumptions and all items loaded on each primary factor (see Table 1). Nevertheless, the reliability indices led us to drop two items, related to increasing structural job resources and increasing social job resources, because of their low item-total correlation.

Confirmatory factor analyses performed on the final 13-item Italian JCS proved the three-factor structure (Model 1), which fitted the data better than the alternative solutions with one factor or two factors.

The reliabilities of the final scales were satisfactory. The Cronbach's alpha coefficients, in particular, are in line with those found by Tims et al. (2012): .81 for increasing structural job resources, .74 for increasing social job resources and .78 for increasing challenging job demands (alphas were respectively .76, .73 and .77 in the original study).

Finally, we investigated the correlations of employees' job crafting behaviors with self-reported work self-efficacy and work engagement, and with performance evaluations expressed by their direct supervisors. The resulting pattern of relations provided additional evidence of the validity of the Italian JCS, in terms of convergent and criterion validity. Indeed, all the three job crafting dimensions were positively associated with the other variables, with the only exception of increasing structural job resources that was not significantly related to job performance. A possible explanation could be that most of the items belonging to this dimension refers to the development of future competences, not directly affecting current goal achievement. Further research is needed to better examine the modest correlations that we found among some of our variables, as reported in details in the Result section. For example, future studies may use a social measure of self-efficacy, which might be more strongly associated with those crafting behaviors oriented toward attaining satisfactory degrees of social interactions or seeking support (i.e., increasing social job resources). All in all, the expected links of job crafting with individual proactivity, operationalized as self-efficacy beliefs, and with desirable individual and organizational outcomes, as employees' engagement and performance, have been supported.

A limitation of the present study is the use of a cross-sectional design that does not permit to establish definite relations of causality between variables. However, the focus was on the validation of the Italian JCS and future longitudinal research can better address patterns of influence between job crafting and other variables. Future studies can also confirm the psychometric characteristics of the instrument on larger samples and considering different classes of employees. Multi-group research design could be useful, for example, to verify potential peculiarities of the construct of job crafting and its dimensions within multiple professional groups. This could contribute to a deeper understanding on how (and whether) diverse types of workers use job crafting strategies differently.

The availability of a tool to measure crafting behaviors can both enhance additional research on the topic and uncover useful practical implications. The questionnaire can

be used, for example, in training or coaching courses aimed at increasing skills of flexibility, initiative and disposition to change. This may provide trainees with an opportunity to check their inclination to job crafting, identifying strengths and areas of improvement related to the forms that job crafting can assume. Moreover, the instrument can be used within the organizational check-up processes, to analyze to what extent job crafting strategies are used and which of these strategies can be promoted to all employees or to specific groups. Finally, the questionnaire can help to recognize job crafting best practices already available in the organization, which may guide social and training activities for newcomers.

References

BALDUCCI, C., FRACCAROLI, F. & SCHAUFELI, W.B. (2010). Psychometric Properties of the Italian Version of the Utrecht Work Engagement Scale (UWES-9). A Cross-Cultural Analysis. European Journal of Psychological Assessment, 26(2), 143-149.

BAKKER, A.B. & DEMEROUTI, E. (2014). Job demands-resources theory. In P.Y. Chen & C.L. Cooper (Eds.), *Work and Wellbeing* (pp. 37-64). Chichester, UK: Wiley-Blackwell.

BAKKER, A.B. & SANZ-VERGEL, A.I. (2013). Weekly work engagement and flourishing: The role of hindrance and challenge job demands. *Journal of Vocational Behavior*, *83*(3), 397-409.

BAKKER, A.B., TIMS, M. & DERKS, D. (2012). Proactive personality and job performance: The role of job crafting and work engagement. *Human Relations*, *65*(10), 1359-137.

BORGOGNI, L., DELLO RUSSO, S., PETITTA, L. & VECCHIONE, M. (2010). Predicting job satisfaction and job performance in a privatized organization. International Public Management Journal, 13(3), 275-296.

CALLEA, A., URBINI, F., INGUSCI, E. & CHIRUMBOLO, A. (2014). The relationship between contract type and job satisfaction in a mediated moderation model: The role of job insecurity and psychological contract violation. *Economic and Industrial Democracy*. Published online before print, DOI: 10.1177/0143831X14546238

PETROU, P. (2013). Crafting the change: The role of job crafting and regulatory focus in adaptation to organizational change. Utrecht University: Utrecht. Available online at http://dspace.library.uu.nl/handle/1874/276440 (last access 11.11.2015)

PETROU, P., DEMEROUTI, E. & HÄFNER, M. (2015). When fit matters more: The effect of regulatory fit on adaptation to change. European *Journal of Work and Organizational Psychology*, *24*(1), 126-142.

PETROU, P., DEMEROUTI, E., PEETERS, M.C., SCHAUFELI, W.B. & HETLAND, J. (2012). Crafting a job on a daily basis: Contextual correlates and the link to work engagement. *Journal of Organizational Behavior*, *33*(8), 1120-1141.

TIMS, M. & BAKKER, A.B. (2010). Job crafting: Towards a new model of individual job redesign. *Journal of Industrial Psychology*, *36*(2), 1-9.

TIMS, M., BAKKER, A.B. & DERKS, D. (2012). Development and validation of the job crafting scale. *Journal of Vocational Behavior*, *80*(1), 173-186.

TIMS, M., BAKKER, A.B. & DERKS, D. (2015) Job crafting and job performance: A longitudinal study. *European Journal of Work and Organizational Psychology*, *24*(6), 914-92

TUCKER, L.R. & LEWIS, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38(1)*, 1-10.

VAN DEN HEUVEL, M., DEMEROUTI, E., BAKKER, A.B. & SCHAUFELI, W.B. (2010). Personal resources and work engagement in the face of change. In J. Houdmont & S. Leka, *Contemporary occupational health psychology: Global perspectives on research and practice* (vol. 1, pp. 124-150). Chichester, UK: Wiley-Blackwell.

ZITO, M., CORTESE, C.G. & COLOMBO, L. (2015). Nurses' exhaustion: the role of flow at work between job demands and job resources. *Journal of nursing management*. Published online before print, **DOI:** 10.1111/jonm.12284

Assessing Primary and Secondary Students' Achievement Goals for Italian and Mathematics Domains: The Italian Version of the Achievement Goal Questionnaire-Revised (AGQ-R)

Daniela Raccanello, Margherita Brondino

Department of Human Sciences, University of Verona, Italy

ABSTRACT. Oltre a rispondere al bisogno di proporre una versione dell'Achievement Goal Questionnaire-Revised (AGQ-R, Elliot & Murayama, 2008) nella lingua italiana supportandone la validità di costrutto e di criterio, questo lavoro ha avuto lo scopo di testarne l'invarianza di misura considerando diversi fattori e di fornire nuovi dati su differenze di età con studenti di scuola primaria e secondaria, di genere e di dominio, nello specifico per estendere la comprensione degli obiettivi di evitamento di prestazione.

SUMMARY. Besides responding to the need to develop a version of the Achievement Goal Questionnaire-Revised (AGQ-R, Elliot & Murayama, 2008) in the Italian language supporting its construct and criterion validity, this work aimed at testing its measurement invariance across a variety of factors and providing new data on cross-sectional age, gender, and domain differences, particularly to extend the understanding of mastery-avoidance goals. The participants were 365 fourth, seventh, and eleventh-graders, who completed two versions of the AGQ-R referred to Italian and mathematics. We also examined responses of the American participants involved in the development of the original instrument. Confirmatory factor analyses supported the goodness of the hypothesized model, characterized by scalar invariance across country, metric invariance across class level, and uniqueness invariance across gender. Structural equation models showed that first-term performance positively predicted the four goal types, while mastery-approach goals positively predicted second-term performance and pleasantness. Achievement goals, higher for Italian for eleventh-graders and females, decreased at increasing ages. Notwithstanding limitations, our data support the validity of this version of the AGQ-R with primary and secondary school students.

Keywords: *Achievement goals, Primary and secondary school students, Native language and mathematics domain*

INTRODUCTION

Understanding motivational processes is one of the core issues of contemporary educational psychology: Such focus mirrors their relevance within learning and teaching contexts, in which mutual interrelations between motivation, cognition, and affect assume a key role for explaining achievement outcomes (Graham & Weiner, 2012). Among different motivational constructs, great attention has recently being paid to achievement goals as "cognitive–dynamic aims that focus on competence" comprising two dimensions: definition in terms of mastery and performance strivings, and valence in terms of positive possibilities to approach success and negative possibilities to avoid failure (Elliot & Murayama, 2008; Hulleman, Schrager, Bodmann & Harackiewicz, 2010).

Referring to definition, back in the eighties the dichotomous achievement goal model distinguished two goal types according to the criteria used for judging competence: reaching competence for mastery goals and focusing on comparisons with others for performance goals (Dweck, 2000). Early empirical findings supported associations between the two goal types and adaptive and maladaptive consequences for learning, respectively, but contradictory results also emerged. More recently, researchers have refined their goal conceptualization paying attention to valence, a second competence-based dimension referring to the ways individuals focus on competence: in terms of associations with positive and desired outcomes for approach goals, or with negative and undesired outcomes for avoidance goals. This distinction was applied first only to performance goals in the trichotomous model and then also to mastery goals in the 2 x 2 achievement goal model, allowing to explain previous inconsistent results concerning mainly performance goals (Elliot & McGregor, 2001). In brief, the 2 x 2 model encompasses four goal types, namely mastery-approach goals and mastery-avoidance goals, "focused on attaining task-based or intrapersonal competence" or "incompetence", respectively, and performance-approach goals and performance-avoidance goals, "focused on attaining normative competence" or "incompetence" (Elliot & Murayama, 2008, p. 613).

As pointed out by a recent meta-analysis including 243 correlational studies for a total of 91,087 participants (Hulleman et al., 2010), predictive validity of valence bifurcation is generally supported, with positive and negative associations with performance and motivational-affective constructs for approach and avoidance goals, respectively. However, caution must be used in interpreting these findings, especially for performance-approach goals for which results are often inconsistent, and for mastery-avoidance goals, only rarely investigated, and for which theoretical and operational issues remain partially undefined (Graham & Weiner, 2012; Hulleman et al., 2010; Payne, Youngcourt & Beaubien, 2007). Moreover, while relationships with performance have been widely investigated, relationships with emotions have only recently being paid attention to, for example documenting that undergraduates' mastery-approach goals positively predicted enjoyment, according to the control-value model of achievement emotions (e.g., Pekrun, Elliot & Maier, 2009). However, relationships between achievement goals and performance are complicated by the moderating role of factors such as nationality, but not class level or gender (Hulleman et al., 2010). For example, comparing American or Canadian versus European samples, correlations with performance are less positive considering mastery-approach goals and more negative considering mastery-avoidance goals for the former compared to the latter, and more negative for Asian versus American or Canadian samples considering performance-avoidance goals. Coherently, in a previous study measuring Italian primary and secondary students' goals with the Patterns of Adaptive Learning Survey (PALS, Midgley et al., 2000), we found that mastery goals positively predicted performance in both native language and mathematics, but no effects of the two performance goals (Authors, 2013).

At present, one of the most used instruments to measure achievement goals (Muis, Winne & Edwards, 2009) is the Achievement Goal Questionnaire (AGQ, Elliot & McGregor, 2001), developed according to the 2 x 2 model. This model represents a viable theoretical framework helping to understand the contributions of goals in educational contexts, as a prerequisite to work on their malleability to improve instructional practice (Graham & Weiner, 2012; Hulleman et al., 2010). Focusing on it as more parsimonious than the more recent 3 x 2 model–in which, besides relying on the approach-avoidance dimension, a further distinction between goals focused on self, task, and other individuals is proposed (Elliot, Murayama & Pekrun, 2011), and therefore involving advantages for young students' comprehensibility, allows to increase our knowledge of mastery-avoidance goals, which so far have received only limited empirical support. The AGQ has been recently revised to solve some conceptual and methodological problems concerning item formulation,

such as reference to constructs like values, concerns, or affect rather than goals; lack of separation between goals and underlying motivations; or absence of content consistency in items focused on different goals. The resulting Achievement Goal Questionnaire-Revised (AGQ-R, Elliot & Murayama, 2008), more rigorous in its correspondence between concepts and their operationalization, revealed good structural validity after being tested with American college undergraduates referring to exam settings. Its predictive validity was also supported, by examining antecedents like need for achievement and fear of failure, and outcomes like intrinsic motivation and performance.

Since its publication, the AGQ-R has been translated into other languages such as Arabic, Greek, or Italian, rarely involving participants younger than university students, and limiting to high school students (e.g., Abd-El-Fattah & Al-Nabhani, 2012; Alkharusi & Aldgafri, 2010; Apostolou, 2013; Authors, 2014). While these studies have documented the goodness of AGQ-R factorial structure, suggesting its generalizability across different nationalities, to our knowledge there is a lack of attention to measurement invariance (except Alkharusi & Aldgafri, 2010, who supported gender invariance with undergraduate Oman students in Arabic, without checking for nationality invariance). Within the translation process of an existing instrument, examining invariance across a variety of factors, and primarily across languages, is essential to demonstrate cultural validity and to make new findings more interpretable, in order to check whether results can be ascribed to group differences or measurement issues (Chen, 2007; Ziegler & Bensch, 2013; Zusho & Clayton, 2011).

Therefore, our aim was to explore some psychometric properties, specifically in terms of construct and criterion validity, of an Italian version of the AGQ-R, whose factorial structure was preliminarily studied with a small sample of university students (Authors, 2014), with primary and secondary school students. To our knowledge, the AGQ-R has rarely been used with these age groups (e.g., Bernacki, Aleven & Nokes-Malach, 2014, involved adolescents), and supporting the validity of its adaptation could help to deepen our understanding of how achievement goals, and particularly the neglected mastery-avoidance goals, are shaped according to factors such as class level, gender, or domain, for which contrasting results exist.

Regarding changes in achievement goals at students' increasing age, many research studies have documented a general decline in the endorsement of mastery-approach goals, performance-approach goals, and performance-avoidance goals, also in the Italian context; however, some authors have reported decreases of mastery goals and increases of performance goals, coherently with the differentiation of student's ability concepts, and others have documented the stability of achievement goal profiles over time (Authors, 2013; Bong, 2009; Dweck, 2000; Paulick, Watermann & Nückles, 2013; Tuominen-Soini, Salmela-Aro & Niemivirta, 2011). Also concerning gender, research has usually neglected mastery-avoidance goals, while a consistent pattern seems to emerge for the other goal types, with females endorsing more frequently mastery-approach goals and males performance-avoidance goals (Gherasim, Butnaru & Mairean, 2013).

Basing on findings documenting both achievement goals' context-specificity when referred to different levels of generality within learning environments (Apostolou, 2013) and the early development of the ability to differentiate motivational beliefs by domain, which gradually refines from school age to adolescence (Bong, 2001), we examined goals separately for two core subjects, native language and mathematics. The two domains differ also for associated stereotypical beliefs, mirroring female superiority for language and male superiority for mathematics (Muzzatti & Agnoli, 2007), thus complicating the influence of gender on the endorsement of achievement goals.

Concerning the construct validity of the AGQ-R, we expected good fit indexes for the model in which the four achievement goals loaded on four separate factors (mastery-approach goals, mastery-avoidance goals, performance-approach goals, performance-avoidance goals), i.e., we expected the items that were designed originally to measure the four achievement goal orientations to load on the four separate factors in the Italian sample (Elliot & McGregor, 2001; Elliot & Murayama, 2008), for both domains. We also tested the structural invariance of the AGQ-R across country (Italy, United States), class level (fourth, seventh, eleventh-graders), and gender (male, female), as a key step in the validation of the instrument for the Italian context and as a way to exclude measurement artefact from subsequent analyses (Ziegler & Bensch, 2013). Moreover, we investigated differences in achievement goals as a function of class level, gender, and domain (Italian, mathematics). We expected scores to be lower at increasing ages (Authors, 2013; Paulick et al., 2013) and higher for Italian compared to mathematics for females, and vice versa for males, mirroring stereotypical beliefs (Muzzatti & Agnoli, 2007); we also explored whether

achievement goal types were endorsed differently, further confirming their differentiation. Concerning criterion validity, for both domains we explored relationships of achievement goals with later school performance and pleasantness, hypothesizing them to be positively predicted by mastery-approach goals, but not by the other goals (Authors, 2013; Hulleman et al., 2010; Pekrun et al., 2009). Finally, we explored whether achievement goals played a partial or total mediating role between first and second-term performance.

METHOD

Participants

The Italian participants were 365 students, including 125 fourth-graders (mean age = 9.85 years, *SD* =.32, range: 9-11 years; 59 female, 66 male), 135 seventh-graders (mean age = 12.98 years, *SD* = .47, range: 12-15 years; 56 female, 79 male), and 105 eleventh-graders (mean age = 16.95 years, *SD* = .41, range: 16-18 years; 64 female, 41 male), nested in 19 classes and 11 schools. They participated on a voluntary basis, following parents' written authorization proposed within the consent form. All the students were guaranteed anonymity, and their teachers were not present while they were answering.

Reference data for testing invariance across country were kindly made available by the authors of the AGQ-R (Elliot & Murayama, 2008). They included the 229 American undergraduates who participated in the original study (mean age = 19.41 years, *SD* = 1.68, range: 17-36 years; 150 female, 76 male, 3 unspecified).

Materials and procedure

We administered a written questionnaire in the classrooms during normal school time, in the second term of the school year in May. We read aloud all the items to avoid missing responses. Each session lasted about 30 minutes. The American Psychological Association ethical standards were followed in the conduct of the study.

– *Achievement Goal Questionnaire-Revised (AGQ-R)*. We proposed two versions of the AGQ-R preliminarily used with Italian university students (Authors, 2014), counterbalanced across classes within each school and class level, referred to two domains, Italian and mathematics (see Appendix; the original items were published in Elliot & Murayama, 2008, p. 617). Each version included 12 items to be evaluated on a 5-point Likert-type scale (1 = *not at all true of me and 5 = very true of me*), presented in the same order as in the original instrument. Three items regarded mastery-approach goals (e.g., "My aim is to completely master the material presented in Italian"), three items mastery-avoidance goals (e.g., "My goal is to avoid learning less than it is possible to learn in mathematics"), three items performance-approach goals (e.g., "I am striving to do well compared to other students in Italian"), and three items performance-avoidance goals (e.g., "My aim is to avoid doing worse than other students in mathematics"). Particular caution was paid to obtain simple linguistic versions for each item, both at the lexical and syntactic level, given that the original version of the questionnaire had been developed to be used with university students (Elliot & Murayama, 2008), and to our knowledge it has been used with adolescents but not with younger students (e.g. Bernacki et al., 2014).
– *School performance*. For each domain, students self-reported their first-term performance (then checked with teachers for reliability) and all of them authorized the school to communicate to us their second-term performance, assigned in June, according to the grades used in the Italian education system (1 = *very low* and 10 = *very high*).
– *Pleasantness*. For each domain, students indicated the level of associated pleasantness on a 10-point Likert type scale (1 = *very low* and 10 = *very high*). Notwithstanding the possible limitations of single-item measures (for example, low variance and reduced validity measuring a complex construct), the literature indicates their reliability and usefulness (Authors, 2013).

RESULTS

We used Mplus version 5.2 (Muthén & Muthén, 1998–2007) to run multilevel confirmatory factor analyses (MCFA), measurement invariance analyses (MI), multilevel structural equation models (MSEM), and path analyses, controlling for a clustering effect of classes in the data. The nested nature of the data (i.e., the fact that the participants belonged to different classes) was taken into account using the Mplus "Complex" syntax, which uses the maximum likelihood estimation with

robust standard errors (MLR) in order to estimate model parameters. We used SPSS version 21.0 for Windows to run all the other analyses.

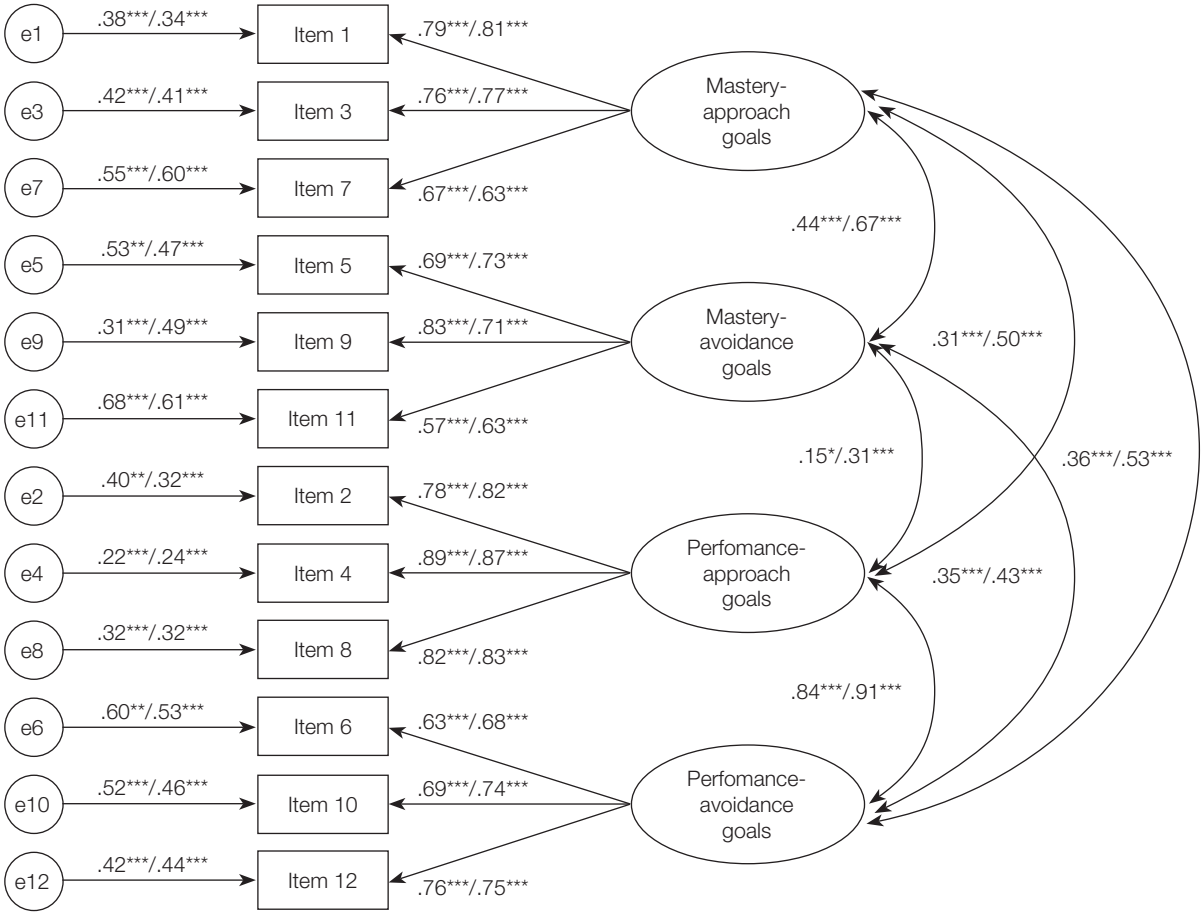
To check for multivariate normality, we verified that skewness (range:.09-.14) and kurtosis (range:.01-1.45) values for each item did not exceed 2.0 and 7.0, respectively, supporting normality assumptions (Curran, West & Finch, 1996). There were no missing data.

Factorial structure of achievement goals

Two multilevel confirmatory factor analyses (MCFAs) separated by domain, clustered by class, allowed to test the goodness of fit of the two hypothesized models, in which the items referred to the four goals load on four distinct latent factors. We considered the Comparative Fit Index (CFI)≥.90, the root-mean-square error of approximation (RMSEA)≤.08, and the Standardized Root Mean Square Residuals (SRMR)≤.11 as threshold values (Beauducel & Wittmann, 2005).

The analyses supported the goodness of the hypothesized models, and verified that the hypothesized factors were measured by single latent variables. In Figure 1, we report the factor models with the parameter estimates. We allowed the four factors to covariate simultaneously. The standardized loadings ranged from .57 to .89 for Italian and from .63 to .87 for mathematics, and they were all statistically significant at the .001 level. The fit indexes for both models were very good (Italian: $\chi^2(48, 365) = 76.46$, $p = .006$; CFI = .98; RMSEA = .04; SRMR = .04; mathematics: $\chi^2(48, 365) = 107.50$, $p<.001$; CFI = .96; RMSEA = .06; SRMR = .05). Therefore, our findings

Figure 1 – Four-factor model for both domains. Read from left to right the digits represent error variances, factor loadings, and latent factor covariances for Italian/mathematics



Note. *$p<.05$; **$p<.01$; ***$p<.001$.

confirmed that the Italian version of the 12 adapted items referred to four distinct latent factors, i.e. mastery-approach goals, mastery-avoidance goals, performance-approach goals, and performance-avoidance goals.

For each domain, responses to items on the same goals were averaged together; all goals intercorrelations were positive (Table 1). We checked for reliability calculating the α-values for each goal; all the α-values were higher than .72, indicating the homogeneity for each construct.

Measurement Invariance (MI)

Measurement invariance (MI) usually relates to how contents of each item are interpreted in the same way across samples (Byrne & Watkins, 2003). If measures of achievement goals operate differently across country, age level, and gender, and these variations are not considered in the measurement, it is inadequate to compare achievement goals or their consequences across groups. For each domain, MI analyses examined hypotheses on the similarity of the covariance structure across groups differing for country, class level, or gender, by considering: (1) configural invariance, allowing all the parameters to be freely estimated; (2) metric invariance, requiring invariant factor loadings; (3) scalar invariance, requiring also invariant intercepts; and (4) uniqueness invariance, requiring invariant item uniqueness. Due to the small and unequal size of our samples, support for noninvariance required Δ CFI≤–.005, supplemented by Δ RMSEA≥.010, for testing metric invariance, and .010 or .005, respectively, for testing scalar and uniqueness invariance (Chen, 2007).

Similarly for the two domains, the base models showed excellent fit indexes for country and gender, but not for class level, presenting poor fit indexes for fourth and seventh-graders. After checking the modification indexes, we directly linked the two mastery-avoidance items focused on the same issue (5, 9) for the three class levels, due to item overlap, and the models improved substantially (Table 2). When we tested simultaneously the different groups not imposing equality

Table 1 –– Intercorrelations, Means (Standard Deviations), and Alpha-values for Scores on Achievement Goals, Performance, and Pleasantness for Italian/Mathematics, respectively

Variable	1	2	3	4	5	6	7	M (SD)
1. Mastery-approach goals	.78/.78	.510***	.409***	.412***	.283***	.401***	.541**	4.08 (.79)
2. Mastery-avoidance goals	.362***	.73/.72	.245***	.331***	.256***	.297***	.344**	3.56 (1.07)
3. Performance-approach goals	.273***	.141***	.86/.88	.750***	.242***	.276***	.373**	3.10 (1.11)
4. Performance-avoidance goals	.267***	.274***	.678***	.73/.76	.202***	.285***	.315**	3.30 (1.08)
5. First-term performance	.321***	.138***	.213***	.148***	–	.815**	.530**	7.14 (1.15)
6. Second-term performance	.285***	.151***	.232***	.190***	.618**	–	.532**	7.25 (1.20)
7. Pleasantness	.490**	.177**	.238**	.153**	.331**	.334***	–	7.62 (1.84)
M (SD)	3.99 (.84)	3.52 (1.04)	3.07 (1.16)	3.27 (1.12)	7.07(1.38)	7.26 (1.36)	6.96 (2.53)	–

Note. Respectively for Italian/mathematics, correlations are presented below/above the diagonal; means (standard deviations) in column/row; alpha-values along the diagonal.
**p<.01; ***p<.001.

Table 2 – Results of Invariance Analyses for the Model of Achievement Goals across Country (Italy, United States), Class Level (Fourth, Seventh, Eleventh-Graders), and Gender (Male, Female), for Italian/Mathematics, respectively

Groups	Model	χ^2	df	*p*	CFI	RMSEA	SRMR	Δ CFI	Δ RMSEA
Country	Italy (n = 365)	76.46/107.51	48	.005/.001	.971/.964	.040/.058	.039/.045	–	–
	United States (n = 229)	78.67	48	.003	.986	.053	.027	–	–
	Configural invariance	217.90/247.55	96	.001/.001	.970/.965	.061/.068	.039/.042	–	–
	Metric invariance	201.43/238.14	104	.001/.001	.975/.967	.056/.066	.052/.062	.005/.002	.005/.002
	Scalar invariance	226.46/281.86	112	.001/.001	.970/.959	.059/.071	.056/.070	.005/.008	.003/.005
	Uniqueness invariance	834.90/866.51	124	.001/.001	.815/.819	.139/.142	.123/.127	.149/.140	.075/.071
Class level	Fourth-graders (n = 125)	54.37/71.25	47	.214/.016	.973/.928	.035/.062	.055/.061	–	–
	Seventh-graders (n = 135)	83.38/87.06	47	.001/.001	.905/.910	.076/.079	.056/.057	–	–
	Eleventh-graders (n = 105)	79.52/77.33	47	.002/.003	.957/.954	.081/.078	.052/.059	–	–
	Configural invariance	272.25/302.28	141	.001/.001	.924/.910	.078/.087	.060/.066	–	–
	Metric invariance	272.24/276.58	157	.001/.001	.924/.926	.078/.079	.072/.068	.000/.016	.000/.008
	Scalar invariance	298.46/317.74	174	.001/.001	.918/.911	.077/.083	.075/.079	.006/.015	.001/.004
	Uniqueness invariance	433.86/590.66	200	.001/.001	.846/.759	.098/.127	.116/.137	.072/.152	.021/.044
Gender	Males (n =186)	82.06/83.13	48	.002/.001	.950/.958	.062/.063	.057/.054	–	–
	Females (n =179)	37.13/71.56	48	.872/.015	1.00/.968	.000/.052	.035/.050	–	–
	Configural invariance	125.68/160.15	96	.073/.001	.983/.964	.034/.054	.049/.053	–	–
	Metric invariance	123.83/158.78	104	.090/.001	.984/.965	.032/.054	.052/.057	.001/.001	.002/.000
	Scalar invariance	131.82/165.66	112	.100/.001	.984/.965	.031/.052	.052/.059	.000/.000	.001/.002
	Uniqueness invariance	145.30/180.63	124	.092/.001	.983/.964	.031/.050	.079/.063	.001/.001	.000/.002

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; Δ CFI = change in Comparative Fit Index; Δ RMSEA = change in Root Mean Square Error of Approximation.

constraints between them (configural invariance), the goodness-of-fit of the models was confirmed across country, class level, and gender. When all factor loadings were constrained to be equal across the three variables (metric invariance), the models resulted invariant for all the three variables; however, for class level (only for mathematics) the Δ RMSEA was below the threshold values, while the Δ CFI was not. When also the intercepts of the observed variables were constrained to be equal across groups (scalar invariance), the models were invariant for country and gender; for class level, the model was invariant for Italian, while for mathematics the Δ RMSEA was below the threshold values but the Δ CFI was not. Finally, when factor loadings, intercepts, and residuals were constrained to be equal (uniqueness invariance), the models were invariant only for gender.

To sum up, the results of the sequence of gradually more restrictive tests of MI supported metric invariance for all the three variables; scalar invariance across country, class level for Italian (and only partially for mathematics), and gender; and uniqueness invariance across gender. In other words, the factorial structure of the models for the two domains was confirmed as substantially invariant, enabling to compare achievement goal levels across the three variables, i.e. country, class level, and gender.

Effects of class level, gender, and domain

A 3 x 2 x 2 x 4 [Class Level (fourth-graders, seventh-graders, eleventh-graders) x (male, female) x Domain (Italian, mathematics) x Achievement Goal Type (mastery-approach goals, mastery-avoidance goals, performance-approach goals, performance-avoidance goals)] repeated-measure ANOVA was carried out on goal scores. Class Level and Gender were treated as between-subjects factors, while Domain and Achievement Goal Type as within-subjects factors. We applied Bonferroni correction to control for Type I error.

This ANOVA revealed main effects of Class Level, $F(2, 359) = 84.71$, $p<.001$, $\eta_p^2 = .32$, and Achievement Goal Type, $F(3, 1077) = 131.81$, $p<.001$, $\eta_p^2 = .27$ (Table 3). Post-hoc *t*-tests indicated that fourth-graders' scores were higher than seventh-graders', in turn higher than eleventh-graders' scores. Concerning goal types, the scores were higher for mastery-approach goals compared to mastery-avoidance goals, higher than performance-avoidance goals, in turn higher than performance-approach goals. Such effects were moderated by a significant Class Level x Achievement Goal Type interaction, $F(6, 1077) = 18.92$, $p<.001$, $\eta_p^2 = .10$, suggesting that the afore mentioned class level differences were less marked for mastery-avoidance goals, for which only fourth and eleventh-graders differed.

In addition, Domain, $F(1, 359) = 6.31$, $p = .012$, $\eta_p^2 = .02$, Domain x Class Level, $F(2, 359) = 7.19$, $p = .001$, $\eta_p^2 = .04$, and Domain x Gender, $F(1, 359) = 8.35$, $p = .004$, $\eta_p^2 = .02$, resulted significant. Scores were higher for Italian ($M = 3.51$, $SD = .72$) compared to mathematics ($M = 3.46$, $SD = .80$). Interpretation of the interactions, confirmed by paired t-tests separated by class level and gender, suggested that it happened only for eleventh-graders [$t(104) = 4.00$, $p<.001$; Italian: $M = 3.01$, $SD = .63$, mathematics: $M = 2.79$, $SD = .61$] and females [$t(178) = 3.97$, $p<.001$; $M = 3.53$, $SD = .72$; $M = 3.38$, $SD = .79$, respectively], with an opposite but not significant trend for males [$t(185) = -1.21$, $p = .228$; $M = 3.49$, $SD = .72$; $M = 3.54$, $SD = .80$].

Relationships of achievement goals with performance and pleasantness

To explore the relationships between achievement goals and performance and pleasantness, we ran multilevel structural equation models (MSEM) for each domain, taking into account the nested nature of the data.

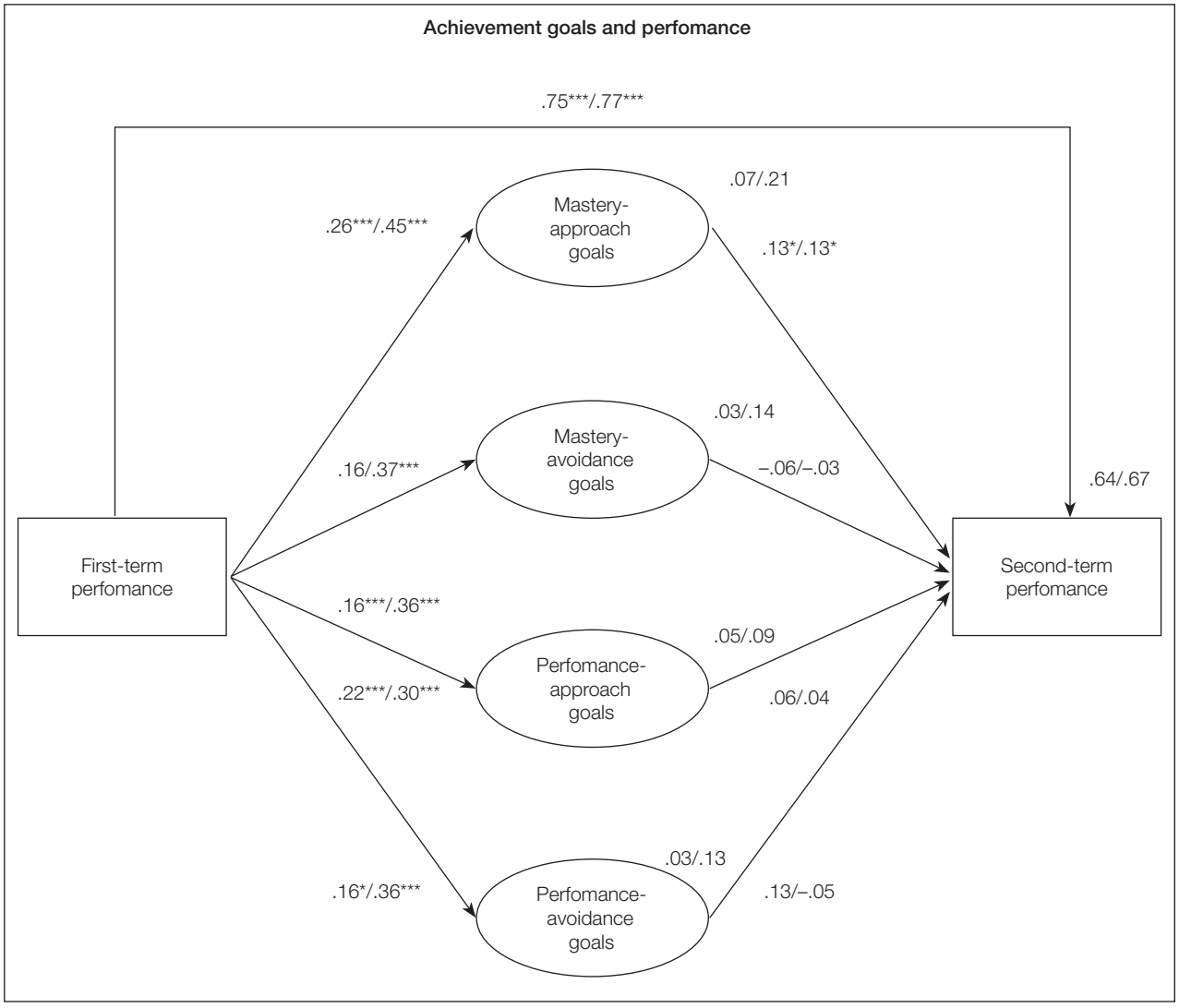
For the first two MSEMs, we considered first-term performance as predictor of achievement goals, and achievement goals as predictors of second-term performance. We also included a direct path between first and second-term performance to verify goals' partial or total mediating role. The models had good fits for both domains [Italian: $\chi^2(64, 365) = 97.19$, $p = .005$; CFI = .98; RMSEA = .04; SRMR = .04; mathematics: $\chi^2(64, 365) = 126.15$, $p<.001$; CFI = .97; RMSEA = .05; SRMR = .04], with significant factor loadings (all $p<.001$) (Figure 2a). All the relationships between first-term performance and goals were statistically significant (except for mastery-avoidance goals for Italian) and positive; only mastery-approach goals positively predicted second-term performance.

It is interesting to note that the effect of first-term performance on achievement goals was stronger for mathematics compared to Italian, as indicated also by the values of explained variances, ranging from .03 to .07 for

Table 3 - Means (Standard Deviations) of Achievement Goals, separately by Class Level, pooled for Domain

Variable	Fourth-graders	Seventh-graders	Eleventh-graders	Total
Mastery-approach goals	4.33 (.56)	4.06 (.73)	3.65 (.66)	4.04 (.71)
Mastery-avoidance goals	3.71 (1.07)	3.54 (.97)	3.35 (.77)	3.54 (.96)
Performance-approach goals	3.68 (.93)	3.22 (.91)	2.19 (.81)	3.08 (1.07)
Performance-avoidance goals	3.85 (.81)	3.45 (.84)	2.41 (.87)	3.29 (1.02)
Total	3.89 (.59)	3.56 (.61)	2.90 (.55)	3.49 (.71)

Figure 2a – Multilevel structural equation models for relationships of achievement goals with performance for Italian/mathematics. Explained variances are reported next to each dependent variable

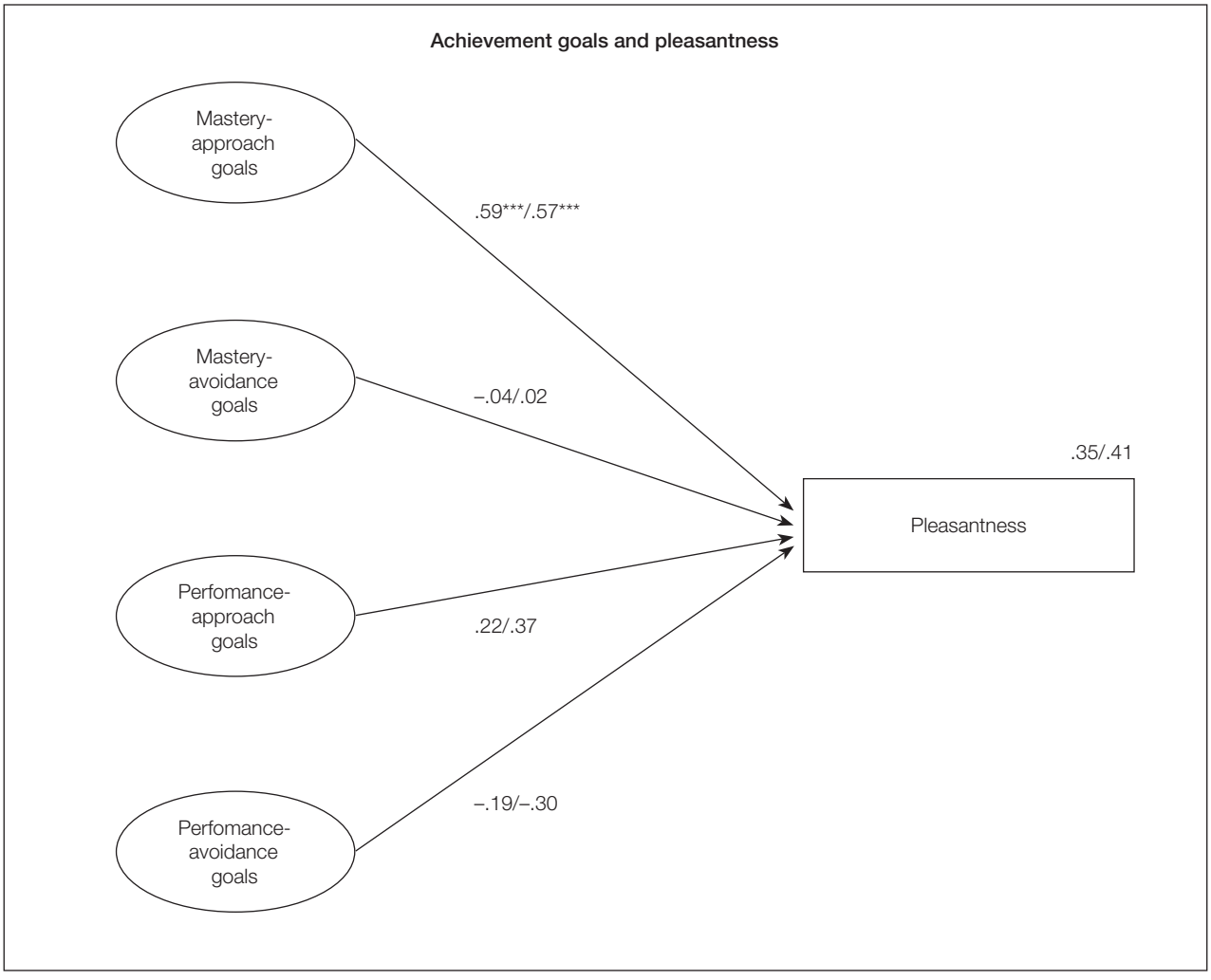


Note. **p*<.05; ****p*<.001.

Italian and from .09 to .21 for mathematics. Furthermore, the direct path between first and second-term performance was statistically significant, highlighting a partial mediating role of mastery-approach goals between first and second-term performance (Italian: indirect effect = .04, p = .017; mathematics: indirect effect = .06, p = .036). In other terms, first-term performance had a strong effect on second-term performance, but this effect was partially due to the influence of students' mastery approach goals on second-term performance. The indirect effect was, again, slightly stronger for mathematics compared to Italian.

For the other two MSEMs, we tested whether achievement goals predicted pleasantness (Figure 2b). The models had good fits for both domains [Italian: $\chi^2(56, 365) = 91.18$, p = .002; CFI = .97; RMSEA = .04, SRMR = .04; mathematics: $\chi^2(56, 365) = 110.97$, p<.001; CFI = .97; RMSEA = .05, SRMR = .04], with significant factor loadings (all p<.001): as expected, mastery-approach goals positively predicted pleasantness. The relationship was slightly stronger for the Italian domain. In general, mastery goals seemed to have a stronger effect for Italian, while performance goals seemed to be more important predictors in mathematics. The explained variance of pleasantness was quite high for both domains (.35 for Italian and .41 for mathematics).

Figure 2b – Multilevel structural equation models for relationships of achievement goals with pleasantness for Italian/mathematics. Explained variances are reported next to each dependent variable



*Note. ***p<.001.*

DISCUSSION

Besides responding to the need to develop a version of the AGQ-R in the Italian language supporting its construct and criterion validity, this work aimed at testing its measurement invariance across a variety of factors and providing new data on cross-sectional age, gender, and domain differences, particularly to extend the understanding of mastery-avoidance goals (Ziegler & Bensch, 2013).

Concerning construct validity issues, we found support of the 2 x 2 hypothesized goal model (Elliot & Murayama, 2008) for Italian and mathematics domains. This suggests the salience of both valence and definition dimensions, at least for the Italian students involved, in the representation of the reasons underlying the endeavours towards learning in specific subjects. Mean differences in the levels with which the students endorsed the four goal types–with the highest scores for mastery-approach goals and the lowest scores for performance-approach goals –further indicated that they are already differentiated in fourth-graders. Even if the higher scores for mastery versus performance goals could be linked to social desirability, the ability to distinguish between the four goal types supports the authenticity of students' responses.

The goodness of the factorial structure of the 2 x 2 model was additionally indicated by the measurement invariance analysis (Chen, 2007; Ziegler & Bensch, 2013). Differently from previous studies in which AGQ-R invariance issues were marginally considered (except Alkharusi & Aldgafri, 2010), we documented the increasing invariance of the model across class level (partially scalar for mathematics and scalar for Italian), country (scalar for both domains), and gender (uniqueness for both domains). Different levels of invariance could depend on limitations of our study such as small sample sizes across class level and American and Italian sample different age. Besides being a prerequisite for the use of the translated questionnaire, these findings allow to exclude the risk of attributing the group differences described to measurement artefacts (Ziegler & Bensch, 2013).

We also reported mean differences in the endorsement of achievement goals according to class level, gender, and domain, as a way to further document how achievement goals are differential in different groups. In support of a detrimental trend for motivation in the transition from primary to secondary school, scores decreased at increasing ages (Authors, 2013; Paulick et al., 2013). This could be linked to a variety of dimensions, such as changes in academic tasks and classroom organization, concurrent psychological development, and changes in peer relationships (Eccles & Roeser, 2011). Concerning domain, we found higher scores for Italian compared to mathematics only for eleventh-graders and females. For older students, this effect could be linked to the increased differentiation between motivational beliefs characterizing them compared to younger students (Bong, 2001). For females, such difference (an opposite though not significant trend characterized males) could mirror stereotypical beliefs on gender superiority in terms of performance in different school domains (Muzzatti & Agnoli, 2007). However, data on superiority derive from tests on cognitive abilities or national surveys and do not correspond to the female advantage in school marks characterizing most subjects from primary to secondary school (Voyer & Voyer, 2014). Making the source of such information salient could be a fruitfully way to diminish the negative consequences associated with gender differences on motivational beliefs (Muzzatti & Agnoli, 2007). Acknowledging differences related to factors such as class level, gender, and domain is relevant not only at a theoretical level, but also from an applied perspective. Knowledge on how students' goals vary according to specific dimensions could be an invaluable instrument for professionals whom daily deals with students, such as teachers, psychologists, or educators. The awareness of these processes is a first step for possible prevention programs aiming at fostering those psychological processes associated with positive performance and wellbeing.

Finally, we examined the causal relationships between goals and performance and pleasantness as a way to provide data on the criterion validity of the AGQ-R. The four goals – besides being strongly correlated with one another, although not sharing neither the valence nor the definition dimension, similarly to data on first to nine-graders (Bong, 2009) – were positively predicted by first-term performance (except for mastery-avoidance goals for Italian, coherently with their focus on avoidance), but only mastery-approach goals positively predicted second-term performance. It seems that past successes or failures in a specific subject influenced the level with which students endorse all types of goals, or, more generally, become motivationally engaged towards a subject. However, only mastery-approach goals play an adaptive role fostering later performance, and their partial mediating role, together with the different signs of coefficients linked to second-term performance, supports goals' construct validity. From a theoretical perspective, these data confirmed

and extended findings characterizing European samples (Hulleman et al., 2010; Pekrun et al., 2009). From an applied perspective, such as clinical or educational, they further stress the need to develop learning environments that can favour students' endorsement of mastery-approach goals (Bong, 2009). Different contextual levels could be taken into account to promote such endorsement, working for example at the individual level on students' awareness of their own goals, at the class level, related for example to the structure of the class goals, or at more comprehensive level, including also contextual goals such as teachers' and parents' goals.

This research study suffers from limitations related for example to the prevailing focus on the 2 x 2 model with respect to other achievement goals frameworks; to the relatively small sample size; and to the use of self-report methods, like desirability effects, or cross-sectional designs, like the absence of control on individual differences. They could be partially addressed in future studies, including for example larger samples to investigate further how class level can moderate associations between goals and outcomes, basing on our preliminary results supporting this effect. However, on the whole our data support the validity of the Italian version of the AGQ-R with primary and secondary school students, making its use worthwhile in learning contexts, as a means to provide new data about the Italian population but having the potentiality to be compared cross-culturally and give innovative contributions to our knowledge on the motivational nuances assuming salience in specific school environments.

Acknowledgements
We wish to thank Andrew Elliot and Kou Murayama who kindly made available the original data of the Achievement Goal Questionnaire-Revised, in order to test invariance across country (Elliot & Murayama, 2008). We also thank Bianca De Bernardi for advice while planning the design, and Claudia Bettin e Manuela Marcolongo for help during data gathering.

References

ABD-EL-FATTAH, S.M. & AL-NABHANI, H.Z. (2012). From self-theories of intelligence to academic delay of gratification: The mediating role of achievement goals. *Australian Journal of Educational & Developmental Psychology*, *12*, 93-107.

ALKHARUSI, H. & ALDGAFRI, S. (2010). Gender differences in the factor structure of the 2 x 2 achievement goal framework. *College Student Journal*, *44*(3), 795-804.

APOSTOLOU, M. (2013). Psychometric properties of the Achievement Goal Questionnaire–Revised administered to Greek university students. *Psychological Reports*, *113*(2), 380-403. doi:10.2466/14.11.PR0.113x22z1

BEAUDUCEL, A. & WITTMANN, W.W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling*, *12*(1), 41-75. doi:10.1207/s15328007sem1201_3

BERNACKI, M., ALEVEN, V. & NOKES-MALACH, T.J. (2014). Stability and change in adolescents' task-specific achievement goals and implications for learning mathematics with intelligent tutors. *Computers in Human Behavior*, *37*, 73-80. doi:10.1016/j.chb.2014.04.009

BONG, M. (2001). Between- and within- domain relations of academic motivation among middle and high school students: Self-efficacy, task-value and achievement goals. *Journal of Educational Psychology*, *93*(1), 23-34. doi:10.1037/0022-0663.93.1.23

BONG, M. (2009). Age-related differences in achievement goal differentiation. J*ournal of Educational Psychology*, *101*(4), 879-896. doi:10.1037/a0015945

BYRNE, B.M. & WATKINS, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, *34*, 155-175.

CHEN, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464-504. doi:10.1080/10705510701301834

CURRAN, P.J., WEST, S.G. & FINCH, J.F. (1996). The robustness of test statistics to non-normality and specification error in confirmatory factory analysis. *Psychological Methods*, *1*(1), 16-29.

DWECK, C.S. (2000). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

ECCLES, J.S. & ROESER, R.W. (2011). Schools as developmental context during adolescence. *Journal of Research on Adolescence, 21*(1), 225-241. doi:10.1111/j.1532-7795.2010.00725.x

ELLIOT, A.J. & McGREGOR, H. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519. doi:10.1037//0022-3514.80.3.501

ELLIOT, A.J. & MURAYAMA, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, *100*(3), 613-628. doi:10.1037/0022-0663.100.3.613

ELLIOT, A.J., MURAYAMA, K. & PEKRUN, R. (2011). A 3 x 2 achievement goal model. *Journal of Educational Psychology*, *103*(3), 632-648. doi:10.1037/a0023952

GHERASIM, L.R., BUTNARU, S. & MAIREAN, C. (2013). Classroom environment, achievement goals and maths performance: Gender differences. *Educational Studies*, *39*(1), 1-12, doi:10.1080/03055698.2012.663480

GRAHAM, S. & WEINER, B. (2012). Motivation: past, present, and future. In K. R. Harris, S. Graham & T. Urdan (Eds.), *APA Educational Psychology, Vol. 1: Theories, Constructs, and Critical Issues*. Washington, DC, US: American Psychological Association.

HULLEMAN, C.S., SCHRAGER, S.M., BODMANN, S.M. & HARACKIEWICZ, J.M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, *136*(3), 422-449. doi:10.1037/a0018947

MIDGLEY, C., MAEHR, M.L., HRUDA, L.Z., ANDERMAN, E., ANDERMAN, L., FREEMAN, K.E., GHEEN, M., KAPLAN, A., KUMAR, R., MIDDLETON, M.J., NELSON, J., ROESER, R. & URDAN, T. (2000). *Manual for the Patterns of Adaptive Learning Scales (PALS)*. Ann Arbor, MI: University of Michigan.

MUIS, K.R., WINNE, P.H. & EDWARDS, O.V. (2009). Modern psychometrics for assessing achievement goal orientation: A Rasch analysis. *British Journal of Educational Psychology*, *79*(3), 547-576. doi:10.1348/000709908X383472

MUTHÉN, L.K. & MUTHÉN, B.O. (1998–2007). *Mplus user's guide* (*5th ed.*). Los Angeles, CA: Authors.

MUZZATTI, B. & AGNOLI, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. *Developmental Psychology*, *43*(3), 747-759. doi:10.1037/0012-1649.43.3.747

PAULICK, I., WATERMANN, R. & NÜCKLES, M. (2013). Achievement goals and school achievement: The transition to different school tracks in secondary school. *Contemporary Educational Psychology*, *38*(1), 75-86. doi:10.1016/j.cedpsych.2012.10.003

PAYNE, S.C., YOUNGCOURT, S.S. & BEAUBIEN, J.M. (2007). A meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology*, *92*(1), 128-150. doi:10.1037/0021-9010.92.1.128

PEKRUN, R., ELLIOT, A.J. & MAIER, M. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology*, *101*(1), 115-135. doi: 10.1037/a0013383

TUOMINEN-SOINI, H., SALMELA-ARO, K. & NIEMIVIRTA, M. (2011). Stability and change in achievement goal orientations: A person-centered approach. *Contemporary Educational Psychology*, *36*(2), 82-100. doi:10.1016/j.cedpsych.2010.08.002

VOYER, D. & VOYER, S.D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174-1204. doi:10.1037/a0036620

ZIEGLER, M. & BENSCH, D. (2013). Lost in translation: Thoughts regarding the translation of existing psychological measures into other languages. *European Journal of Psychological Assessment*, *29*(2), 81-83. doi:10.1027/1015-5759/a000167

ZUSHO, A. & CLAYTON, K. (2011). Culturalizing achievement goal theory and research. *Educational Psychologist*, *46*(4), 239-260. doi:10.1080/00461520.2011.614526

APPENDIX

Italian Achievement Goal Questionnaire-Revised (AGQ-R) Items (adapted from Elliot & Murayama, 2008)

Item	Item content
	Mastery-approach goals
1	Il mio scopo è padroneggiare pienamente gli argomenti spiegati in italiano/matematica.
3	Il mio obiettivo è imparare il più possibile in italiano/matematica.
7	Mi sforzo di capire i contenuti dell'italiano/della matematica nel modo più completo possibile.
	Mastery-avoidance goals
5	Il mio scopo è evitare di imparare meno di quanto potrei in italiano/matematica.
9	Il mio obiettivo è evitare di imparare meno di quanto sia possibile imparare in italiano/matematica.
11	Mi sforzo di evitare una comprensione incompleta degli argomenti dell'italiano/della matematica.
	Performance-approach goals
2	Mi sforzo di andare bene in confronto agli altri studenti in italiano/matematica.
4	Il mio scopo è ottenere buoni risultati rispetto agli altri studenti in italiano/matematica.
8	Il mio obiettivo è riuscire meglio degli altri studenti in italiano/matematica.
	Performance-avoidance goals
6	Il mio obiettivo è evitare di ottenere risultati scarsi in confronto agli altri in italiano/matematica.
10	Mi sforzo di evitare di riuscire peggio degli altri in italiano/matematica.
12	Il mio scopo è evitare di andare peggio degli altri studenti in italiano/matematica.

Premises for innovation: Italian validation and dimensionality of the Inventory of Organizational Innovativeness (IOI)

Maria Luisa Farnese[1], Roberta Fida[2]

[1] Department of Psychology – Sapienza University of Rome
[2] Norwich Business School - University of East Anglia

ABSTRACT. Il lavoro si propone di contribuire alla validazione dell'IOI-Inventory of Organizational Innovativeness (Tang, 1999), uno dei pochi strumenti multidimensionali per la misura dell'orientamento delle organizzazioni verso l'innovazione. I risultati di uno studio su 616 lavoratori italiani non confermano la struttura teorica a 9 fattori ma confermano quella a 6 fattori emersa in altri studi. Le analisi delle proprietà psicometriche della scala e delle sue relazioni con costrutti affini confermano la validità e affidabilità dell'IOI per rilevare i diversi aspetti che contribuiscono a promuovere la capacità di una organizzazione di essere innovativa.

SUMMARY. Literature underlines the role of the organizational orientation toward innovation as a precursor of its effective capability to generate and adopt innovations, in this way gaining competitive advantages. However less attention has been devoted to the methodological issues concerning how to measure this construct. Indeed, the few existing measures are often one-dimensional and neglect the multiple facets of this construct. In this paper we examine the multidimensional IOI-Inventory of Organizational Innovativeness (Tang, 1999) with the aim of verifying its psychometrics properties, validating it in the Italian context, and exploring the relationships among its dimensions and other related constructs (servant leadership, climate for support to innovation, climate for participative safety) and outcomes (performance and innovation adoption). Results of the confirmatory factor analysis on a sample of 616 Italian employees did not support the theoretical 9-factor structure. The subsequent exploratory factor analysis attested for a 6-factor model in line with the empirical solution emerged in a previous research. Results of the correlations confirmed the relationship of the IOI's dimensions with both correlated and outcomes measures. Overall, findings of this study attested for the good psychometric properties of the IOI and support that this inventory is a reliable and valid measure of the organizational orientation toward innovation to be used to assess the different facets that contribute to promote the innovation adoption.

Keywords: *IOI; Organizational innovativeness; Orientation to innovate; Innovation adoption; Inventory*

INTRODUCTION

In the present age of rapid change, innovativeness is the main resource which allows organizations to face the increasing and unstable demands from their environment and to gain competitive advantages. Many scholars have demonstrated, in line with the seminal contribution of Zaltman, Duncan & Holbeck on innovation (1973), that the orientation toward innovation is an important precursor of the concrete innovation implementation stage (Berthon, Hulbert & Pitt, 1999; Hurley & Hult, 1998; Hurley, Hult & Knight, 2005), organisation's performance and economic growth (see Siguaw, Simpson & Enz, 2006). Within this approach, orientation to innovation expresses the degree to which the members of an organization are willing or not to consider the adoption and are committed to their use, as well as the degree to which the management recognizes and takes care of the need for new ideas and actions (Van de Ven, 1986).

From a managerial point of view, orientation toward innovation has been conceived as a strategic competitive orientation (Lynch, Walsh & Harrington, 2010; Manu, 1992) and a key organisational resource (Menguc & Auh, 2002). Indeed, "a firm's long-term success may rely more on an overall firm-level innovation orientation that produces capabilities that spawn innovations, and less on specific innovations" (Siguaw et al., 2006, p. 557). However, research has primarily focused on product and process innovations and on structural factors affecting the innovation outputs (Simpson, Siguaw & Enz, 2006), leaving quite unexplored the role of orientation to innovation in enhancing innovation (Cepeda-Carrion, Cegarra-Navarro & Jimenez-Jimenez, 2011; Hurley & Hult, 1998; Tang, 1999).

In addition, although the different definitions of orientation to innovation highlight the multidimensional nature of this construct (e.g. Amabile, 1997; Lynch et al., 2010), it has been commonly operationalized as one-dimensional, and most empirical evidences are based on the scale developed by Hurley and Hult (1998; e.g. Calantone, Garcia & Droge, 2003; Cepeda-Carrion et al., 2011; Zhou, Gao & Yang, 2005). To the best of our knowledge, only the IOI–Inventory of Organizational Innovativeness developed by Tang (1998) captured the orientation to innovation multi-dimensional nature. In fact, some other scales only pick some of the different facets related to the orientation to innovation, including dimensions referred to both innovative orientation and innovation outcomes (such as the Wang and Ahmed's (2004) questionnaire).

The purpose of this paper is to analyse the dimensionality and reliability of the IOI (Tang, 1998) within the Italian context. Moreover it aims to contribute to the IOI construct validation by examining the relationships of its dimensions with other constructs related to the organisational orientation toward innovation (i.e. servant leadership, climate for support to innovation, climate for participative safety; Hulsheger, Anderson & Salgado, 2009; West & Anderson, 1996) as well as with some organisational outcomes (i.e. organizational performance and innovation adoption; Paleo & Wijnberg, 2008; Parris & Peachey, 2013; Tang, 1999; van Dierendonck, 2011).

THEORETICAL FRAMEWORK

The orientation toward innovation dimensionality

While some authors defined orientation toward innovation as a unitary construct (Hult, Hurley & Knight, 2004), most of them highlighted the different facets of it. For instance, Amabile (1997) asserted that "the most important elements of the innovation orientation are: the value placed on creativity and innovation in general, an orientation toward risk (versus an orientation toward maintaining the status quo), a sense of pride in the organization's members and enthusiasm about what they are capable of doing, and an offensive strategy of taking the lead toward the future (versus a defensive strategy of simply wanting to protect the organization's past position)" (p. 52). Similarly, in their literature review, Lynch and colleagues (2010) conceptualized innovativeness as a multidimensional construct which includes five key components: *creativity*, or the firm's capability to produce new and distinctive ideas, exceeding routine; *openness to new ideas*, or receptiveness to and tolerance of new ideas and experiences; *intention to innovate* (strategic willingness, commitment to innovate); *willingness for risk-taking*, or coping with uncertainty and ambiguity connected to innovation; and *capacity to innovate*, or the necessary skills, knowledge, capabilities and other distinctive resources readily available to adopt or implement new ideas or to take advantage of market opportunities.

Tang (1998), as well, conceptualized orientation toward innovation as a multifaceted construct, assuming a dynamic perspective that simultaneously includes nine dimensions

concerning different organizational levels. Some of them are related to the general organizational commitment toward innovation: *management* support, giving adequate resources and adopting coherent opportunities and rewards to promote innovation; *raising projects*, that is being active in collecting ideas, making suggestions, and exploring through new projects; *doing projects*, expressing the capability to organize, clearly define, implement, and monitor projects; *information and communication* processes that allow dissemination of relevant information, access to documentation and database, and the capturing of ideas and opportunities both from internal and external sources. Other dimensions refer to the interpersonal level: the degree supervisors adopt a consultative and flexible *leadership* style; the degree of teamwork *integration* and mutual trust, being capable to work together harmoniously; the degree colleagues adopt supporting and helpful *behaviour* for work. The last dimensions are related to the job level. One describes the degree colleagues have k*nowledge and skills* useful to generate new ideas and create intellectual assets and to turn ideas into action. The other the degree they carry out intellectually stimulating, non-routine and challenging *tasks* that allow creativity and exploration.

The author (Tang, 1999) operationalized the construct in the *IOI-Inventory of Organizational Innovativeness*, a 44-item scale aimed to measure the aforementioned nine facets composing the organizational orientation toward innovation. However, the empirical study he carried out in a professional engineering society did not confirm the nine-factor structure. Specifically, it resulted in a 6-factor empirical solution. Although the authors did not published this last solution, so it is not clear which items loaded in which facets[1] , the first factor captures the organisation capability of *doing projects*; the second one mainly captures both the managerial *support* and the *information and communication* theoretical dimensions; the third one captures both the *leadership* and the *raising projects* theoretical dimensions; the fourth factor captures mainly the *tasks* dimension; the fifth factor captures both the *behaviour* and the *integration* theoretical dimensions; and the last factor captures the *knowledge and skills* theoretical dimension. To the best of our knowledge only few authors have used the IOI. For instance, Aliaga (2005) used it taking for granted the nine factors and proposing a revised version of the inventory. Other authors used part of the IOI, selecting some dimensions or items and showing their relationship with organizational product and process innovation (see for example, Prajogo & Ahmed, 2006; Prajogo, Power & Sohal, 2004).

The first aim of the present study is to examine the IOI psychometric properties, contributing to its cross-cultural validation and generalization. Specifically, according with the conceptualisation of the IOI (Tang, 1998), we will first examine the 9-factor structure, then we will examine the reliability of each of the facets, and finally we will examine the relationships of the IOI facets with relevant correlates and outcomes.

Relationship between orientation toward innovation and related constructs

Literature identifies many "soft" factors enabling the innovation implementation. For instance, leadership has been proved to be a trigger of organizational and individual innovation (de Jong & den Hartog, 2007). Specifically, servant leaders, by focusing on employees' empowerment, enhance their values and abilities, encourage participation in decision-making and information sharing, and coach them for innovative performance (Konczak, Stelly & Trusty, 2000). In addition, by promoting a safety climate, they lead to interpersonal acceptance, reduce power distance, and so in doing to learning. In this way, servant leaders foster employees' attachment to the organization, cooperative and extra-role behaviours, and promote a higher engagement in challenging tasks, so that these will results in organizational effectiveness and willingness to change (Parris & Peachey, 2013; van Dierendonck, 2011).

Other dimensions that have been identified as significant antecedents of effective innovation implementation are those related to the climate for innovation, which can be defined as the perception that involvement in innovation is widespread among group members (i.e. Carter & West, 1998; Ekvall, 1996). Specifically, the team climate dimension of *support for innovation* expresses the degree of support (i.e. available resources and time, cooperation, practical support) that teammates feel they receive to enhance the generation

1 For what we know, detailed results of the explorative factor analysis have not been published, so it is not clear which items load in which factors, but only how many items of each dimension load on each factor. Anyhow Tang (1999), for further analyses (e.g. the overall profile, differences in effectiveness and innovativeness performance, comparison between managers and non-managers), doesn't take into account this empirical solution and refers to the theoretical dimensions.

and the development of organizational innovative processes (Anderson & West, 1998). So, employees feeling this climate perceive the innovation as a collectivistic process, depending on the commitment of the whole group, that cooperate, share responsibilities and help if needed (Anderson & West, 1998). Similarly, the *participation safety* of team climate dimension expresses the perception of non-threatening and not-judging interpersonal relationships, which increases teammates' interaction and motivates their participation in decision-making and information sharing processes. Hence, it encourages to express divergent ideas and to improve ways of working. Overall, both these climate dimensions lead to organizational innovative outcomes (Adams, Bessant & Phelps, 2006; Bain, Mann, Pirola-Merlo, 2001; Curral, Forrester, Dawson & West, 2001; Hulsheger et al., 2009). Specifically, support for innovation proved to be a predictor of overall innovation and of the number of innovation novelty, and participation safety resulted to be the best predictor of the number of innovations (West & Anderson, 1996).

As already mentioned, in order to further investigate the construct validity of the IOI questionnaire, the second aim of this study is to examine the relationship among the IOI dimensions and both servant leadership and the two climate for innovation dimensions (support for innovation and participative safety). Specifically we hypothesize that all of them will be positively related to each of the IOI dimensions.

Many scholars have also highlighted that the orientation to innovation influences the effective generation and adoption of new products/services (Paleo & Wijnberg, 2008; Prajogo & Ahmed, 2006; Prajogo et al., 2004; Zaltman et al., 1973) and so it affects organisation's concrete innovative capacity (Woodside, 2005). Despite this strong theoretical framework, only few scholars provided empirical evidence of the relationship between organizational innovativeness and performance (Hult et al., 2004) or innovative capacity (Hurley & Hult, 1998). Thus, another aim of this paper is to fill this gap and examine whether and how the IOI dimensions will be positively related to both organisational performance and innovation adoption.

METHOD

Participants and procedure

Participants of this study were 616 Italian employees working in different sectors (see Table 1). Respondents were mainly males (62%), ranged in age from 19 to 62 years (M = 44 years, SD = 2.54) and had attained a relatively high level of education (41% graduate, 47% high school). Participants held different organizational positions (42% operatives, 39% technical-specialized, 19% management) and ranged in organizational tenure from 1 to 38 years (M = 14 years, SD = 13.6).

Data collection was conducted by research assistants. Specifically each of them directly contacted the company's managers and after their approval they administered the questionnaire. Participants voluntarily participated in the study and did not receive any kind of reward. Each of the employees received the questionnaire in a blank envelope and a presentation letter, which contained a brief description of the research and its main objectives. Prior to administering

Table 1 – Productive sectors of the sample

Productive sectors	n	%
Aviation industry	100	16.2
Pharmaceutical industry	60	9.7
Insurance	202	33.1
Marketing	100	16.2
Consulting and development	100	16.2
Railways	12	1.9
Public health	42	6.8
Tot.	616	100%

the surveys, all participants were informed of the anonymity and confidentiality of the survey and were allowed to decline participation if they so choose. To ensure heterogeneity of the sample, each research assistant approached between 10 and 30 employees from different organisations (see Table 1).

Measures

Orientation toward innovation. The organizational orientation to innovation was measured with the IOI-Inventory of Organizational Innovativeness developed by Tang (1999). This questionnaire, already described in the introduction section, is a 44 item self-report scale and it has been developed for measuring 9 theoretical dimensions. Participants were asked to rate their level of agreement for each of the item on a five-point Likert scale (from 1 = s*trongly disagree* to 5 = *strongly agree*). Two bilingual researchers independently translated the original scale into Italian, then a discussion was followed in order to produce the final Italian version (see Appendix).

Servant Leadership was measured by adapting the 6-item scale by Ashill, Carruthers & Krisjanous (2006). It assesses leader's active engagement in helping and meeting the employees' needs, and his/her role in creating an environment conducive to high quality products-service. Example items are: "Management regularly spends time 'on the floor' (with clients and frontline staff)", "Management provides resources, not just 'lip service', to enhance my ability to provide excellence products-service". Employees expressed their degree of agreement on a 5-point Likert scale, from 1 = *strongly disagree* to 5 = *strongly agree* (Cronbach's alpha = .87).

Team climate for innovation: both *Support for innovation and Participation safety* scales were measured with 16 items from TCI–Team Climate Inventory, developed by Anderson and West (1998). The first dimension (7 items) assesses the perception of support and resources given by teammates to other members of the group for the development of new ideas or to solve problems (e.g. "In this team we take the time needed to develop new ideas"; "Members of the team provide and share resources to help in the application of new ideas"). *Participation safety* (9 items) assesses the perception of trusty, not-threatening interpersonal relationships, so that teammates feel they can safely offer new ideas, share information and participate in decision-making (e.g. "Everyone's view is listened to even if it is in a minority", "We share information generally in the team, rather than keeping it to ourselves"). Respondents were asked to indicate the extent to which each statement was true for their team on a 5-point Likert scale, from 1 = *strongly disagree* to 5 = *strongly agree* (both Cronbach's alphas = .93).

Perceived organizational outcomes. As indicators of the overall organizational performance, we used the two self-report assessment items included in the IOI (Tang, 1999):"My organization is effective in innovating" and "Overall, my organization is an effective organization". For these two items we asked participants to rate the level of agreement on a 5 point Likert scale (from 1 = *strongly disagree* to 5 = *strongly agree*). In addition, we measured innovation adoption with two items developed for the scope of this research. Specifically, we asked to participants whether, in the last three years, their organization has introduced into the market new products or services ("We placed new products on the market"; "We proposed new services for our customers"). For each of these items participants indicated the frequency of innovations, using a 5-point Likert scale (from 1 = *never* to 5 = *often*).

Data analysis

To validate the I*OI-Inventory of Organizational Innovativeness*, its psychometric properties were investigated. In particular, Confirmatory Factor Analysis (CFA) was used to test the theoretical IOI 9-factor structure. The model fit was analysed by examining along with the chi square, the Comparative fit index (CFI), the Tucker-Lewis Index (TLI); root mean square error of approximation (RMSEA) and standardised root mean square residual (SRMR) (Byrne, 2012; Meade, Johnson & Braddy, 2008).

After ascertained the dimensionality of the scale, the reliability of each dimension was analysed. Specifically, Cronbach's Alpha, Composite Reliability (CR) and the Maximal Reliability (MR) (see Fornell & Larcker, 1981; Raykov & Marcoulides, 2011) were examined. For these coefficients, values approaching 1 support the good reliability of the measure assessing the underlying latent construct (Raykov & Marcoulides, 2011). Furthermore, construct validity was examined by correlating the IOI dimensions with different types of correlates. Specifically we examined the association with leadership and team climate for innovation and with some outcomes related to innovation adoption and organizational effectiveness. Data were analysed by using Spss and Mplus softwares.

RESULTS

Before proceeding with the analysis, the normality of all the items of the scale was ascertained. Specifically, skeweness and kurtosis indices ranged from .001 to .931. Given that all items were normally distributed, EFA was performed using Maximum Likelihood parameter estimates.

Psychometric properties of the Tang's IOI– Inventory of Organizational Innovativeness

Results of the CFA attested for a not satisfactory fit ($\chi^2_{(866)}$ = 3481.62; p<.001; RMSEA = .070 (.068–.073); p<.01; CFI = .87; TLI = .86; SRMR = .06). In addition, the analysis of the correlations among the IOI facets showed that some dimensions were highly correlated each other (i.e., *Raising projects with management Support* .88; *Doing projects with Information and Communication* .87; *Raising projects with Integration* .82). These high correlations between some facets suggested for a more parsimonious solution, in line with the empirical study by Tang (1999) in which the author extracted 6 factors. Hence, we decided to test an exploratory factor analysis extracting as suggested 6 factors.

Results of the six-factor model showed the following fit indices: $\chi^2_{(697)}$= 2637.025; p<.001; RMSEA = .067 (.065–.070); SRMR = .029. Since this solution revealed eight items which loaded lower than |.30| or with higher loadings in more than one factors, they were deleted, and a second analysis was performed. The final six-factor model (Table 2) fits the data – $\chi^2_{(429)}$ = 1760.81; p<.001; RMSEA = .071 (.068–.075); SRMR = .028 – explaining 62.4% of the total variances.

The first three factors gather, each, two theoretical dimensions, while the other three factor correspond each one to the IOI theoretical dimensions. The first factor accounted for 14.5% of the total variance and gathered items from *Doing projects* and *Information and Communication* dimensions. It was labelled *Alignment* since it is related to the organization's capability to manage a project clearly defining its goals, monitoring and evaluating it, giving adequate resources, and ensuring that all those in need have access to documentation and databases and all relevant information. Thus, it is related to the organization's orientation to enhance the coherence between objectives and resources and among all actors involved, so that an innovation may be implemented and become effective.

The second factor accounted for 15.5% of the total variance and gathered items from *Raising projects* and management *Support* dimensions. It was labelled *Promotion* since it relates to the organization's openness toward the generation of new ideas, suggestions and project proposed by employees for work improvement and innovation. This implies both a psychological safety climate and processes that enhance and support raising ideas through opportunities and reward systems.

The third factor accounted for 13% of the total variance and gathered items from *Behavior* and *Integration* dimensions. It was labelled *Team support* and refers to the degree of integration and perceived support from colleagues that are considered helpful, trustworthy, willing to cooperate, thereby enhancing teamwork and cohesion.

The fourth factor accounted for 8% of the total variance and included the items of the *Leadership* theoretical dimension. It refers to the perception of top managers as available to listening and communicating, capable of adopting a consultative style, valuing employees' opinions and motivating them towards innovation and work improvement. Thus, it is a factor that expresses the general organization's commitment toward innovation through human resources' motivation and direction.

The fifth factor accounted for 7% of the total variance and was labelled Task since it involves items of this theoretical dimension. It expresses the degree to which employees consider their work intellectually stimulating and challenging, based on creativity and on the capability to manage non-routine issues, work that gives them the opportunity to learn and to explore.

And finally, the sixth factor accounted for 5% of the total variance and is composed of the items of the *Knowledge & skills* theoretical dimension. It refers to the perception of teammates as a resource for organizational life and development because of the strength of their knowledge, skills and creativity, and because of their ability to implement new ideas. As shown in Table 3, factors were correlated and ranged from .36 (correlation between Knowledge & skills and Task) to .67 (correlation between Promotion and Alignment).

Table 4 reported the Cronbach's Alpha, maximal and composite reliability, corrected item-scale correlations range, and standard deviation for each factor. As shown, all IOI dimensions were reliable.

Table 2 – Final version of the exploratory factor analysis of IOI– Inventory of Organizational Innovativeness

		Factors					
Item number	Dimension	Alignment	Promotion	Team support	Leadership	Task3	Knowledge & skills
33	DOP–3	**.86**	.01	–.02	–.08	.05	.00
34	DOP–4	**.75**	.06	–.04	.00	.03	.03
32	DOP–2	**.74**	.14	.08	–.04	–.03	–.11
35	DOP–5	**.68**	–.07	.05	.18	–.01	–.04
42	IC–2	**.62**	.10	.01	.08	–.10	.05
31	DOP–1	**.56**	.13	.19	.14	–.07	–.10
43	IC–3	**.53**	.20	–.05	–.09	.08	.07
41	IC–1	**.50**	.15	.21	.14	–.12	–.04
44	IC–4	**.48**	.12	.18	–.07	.08	.06
27	RAP–3	.12	**.73**	.07	–.05	.08	–.06
8	SUPP–4	.07	**.68**	–.10	.10	–.03	.15
9	SUPP–5	.21	**.63**	–.17	.08	–.03	.08
30	RAP–6	.01	**.62**	.05	–.06	.11	–.09
7	SUPP–3	.13	**.61**	–.09	.16	.03	.08
6	SUPP–2	.07	**.60**	–.05	.15	.06	.10
26	RAP–2	.07	**.54**	.27	–.10	.09	.06
25	RAP–1	.17	**.54**	.26	–.07	.08	–.06
28	RAP–4	.13	**.52**	.12	.03	.14	–.10
19	BEH–3	.01	–.18	**.73**	.07	.11	.06
23	INT–3	.00	.17	**.71**	.08	–.12	–.02
22	INT–2	.02	.26	**.69**	.01	–.14	–.03
17	BEH–1	.08	–.23	**.68**	–.05	.09	.18
18	BEH–2	.03	–.02	**.66**	.01	.09	.03
21	INT–1	.05	.14	**.65**	–.07	.07	–.02
24	INT–4	.22	.15	**.42**	.11	–.02	.06
1	LEA–1	.05	–.02	.03	**.80**	.01	.02
2	LEA–2	.10	.12	–.03	**.72**	–.01	.05
3	LEA–3	.24	–.05	–.05	**.66**	.15	.03
4	LEA–4	–.14	.27	.21	**.65**	.01	–.13
14	TASK–3	–.08	.12	–.07	–.01	**.81**	.07
16	TASK–5	.12	.05	.01	.02	**.71**	–.05
15	TASK–4	–.14	.08	.15	.04	**.63**	–.06
12	TASK–1	.16	.02	.09	.10	**.54**	.04
39	KNS–4	–.07	.07	.09	–.00	.02	**.85**
38	KNS–3	.06	–.04	.24	.01	–.05	**.57**
40	KNS–5	.04	.27	.19	.02	.05	**.43**

Note: Items refer to the following theoretical dimensions: DOP = doing projects; IC = information and communication; RAP = raising projects; SUPP = support; TASK = task; BEH = behaviour; INT = integration; LEA = leadership; KNS = knowledge and skills;.

Table 3 – Correlations among the IOI's dimensions

	1	2	3	4	5
1. Alignment	–				
2. Promotion	.67**	–			
3. Team support	.60**	.55**	–		
4. Leadership	.57**	.57**	.47**	–	
5. Task	.38**	.53**	.52**	.37**	–
6. Knowledge & skills	.48**	.37**	.49**	.38**	.36**

Note. $^{**}p < .01$

Table 4 – IOI reliability coefficients

	Cronbach's Alpha	Maximal reliability	Composite reliability	Corrected item-scale correlations
1. Alignment	.92	.92	.93	.64–.76 (M = .72, SD = .04)
2. Promotion	.93	.93	.93	.60–.77 (M = .74, SD = .06)
3. Team support	.90	.90	.90	.67–.73 (M = .70, SD = .02)
4. Leadership	.81	.90	.90	.73–.79 (M = .77, SD = .02)
5. Task	.88	.85	.84	.63–.74 (M = .71, SD = .04)
6. Knowledge & skills	.81	.81	.83	.60–.76 (M = .66, SD = .08)

Relations among IOI dimensions and other constructs

Table 5 reported the analysis of the correlations among the six IOI dimensions, servant leadership, climate for innovation dimensions (support for innovation and participative safety), organisational performance and innovation adoption. As hypothesized, all six IOI dimensions showed a significant positive correlation with these dimensions.

DISCUSSION

The main aim of the present study was to investigate the factorial validity of the Italian version of the IOI– Inventory of Organizational Innovativeness by Tang (1999), one of the few instruments adopting a multidimensional perspective to operationalize the construct. Results of the present study support the good psychometric properties of the IOI questionnaire in the Italian context, although partially confirming the theoretical structure proposed by the author. Indeed, the CFA did not provide support for a nine-dimensions scale. A further EFA suggested for a six-factor structure in which items of different dimensions loaded together in a single factor. Findings showed that eight items had poor relationship with the latent dimensions, thus have been dropped from the analysis.

Specifically, the first IOI factor, *Alignment* (gathering items from the *Doing projects* and *Information* and *Communication dimensions*), allows to describe organizations in relation to their orientation toward a shared planning, with clear strategies and objectives, where everyone receives adequate resources and access to relevant information. The second factor, *Promotion* (gathering items from the *Raising projects* and management *Support* dimensions), allows to describe organizations capable to raise projects and to give support to innovation: it expresses the willingness to accept new ideas and projects, recognize and reward innovative employees, and concretely give resources and opportunities to enhance innovation. These two dimensions are mainly referred to the organizational level, depicting the organizational strategic orientation toward innovation, that is pursuing innovation as a main goal and providing its members with the necessary support and conditions so that it can be implemented.

Table 5 – Correlations among the IOI's dimensions and other variables

	Servant leadership	Team climate for innovation		IOI's organizational performance assessment		Innovation adoption	
		Support for innovation	Participati-ve safety	Effectiveness in innovating	General effectiveness	Product innovation	Service innovation
1. Alignment	.64**	.61**	.56**	.70**	.77**	.30**	.41**
2. Promotion	.57**	.66**	.51**	.65**	.64**	.26**	.32**
3. Team support	.51**	.64**	.68**	.55**	.65**	.34**	.40**
4. Leadership	.63**	.52**	.51**	.46**	.61**	.26**	.35**
5. Task	.40**	.55**	.46**	.49**	.47**	.27**	.31**
6. Knowledge & skills	.43**	.58**	.55**	.53**	.57**	.32**	.33**

Note. $^{**}p < .01$

The *Team support* dimension (gathered items from the *Behavior* and *Integration* dimensions) allows describing organizations in which colleagues are willing to share knowledge and to take initiatives, are helpful and supportive for work and difficulties and, overall, trustworthy. The *Leadership* dimension is related to the adoption of a consultative and flexible style, aimed at empowering human resources and challenging them to work improvement. These dimensions are related to the interpersonal level, expressing the degree the vertical and peer relationships are perceived as supportive for the knowledge and ideas sharing and the teammates' commitment to innovation.

The *Knowledge & skills* dimension is related to teammates' perception as strong in knowledge and abilities, capable to implement innovative ideas. The last dimension, *Task*, allows to describe organisations promoting a non-routine, creative and challenging work. These latter dimensions are referred to the individuals' level, highlighting the organizational orientation to invest on its members and their professionalism as a resource for innovation.

Currently, to our knowledge, no studies have investigated the factorial structure of the IOI, also in other national contexts, thus this result also provides a more general empirical support to the IOI validation.

In addition, the six dimensions provide a composite picture about how different components of the organization can contribute to promote its innovativeness. Consistently with our hypotheses, the analysis of the correlations showed positive relations among all the IOI dimensions and the other cultural factors for innovation we considered. Specifically, organizations high in IOI dimensions related to the organizational (Alignment, Promotion) and interpersonal (Leadership, Team support) levels, also perceive their leaders as adopting a servant style, that is supportive for the employees' motivation, participation and direction. Furthermore, their members feel a team climate enhancing the teammates cooperation to implement new ideas and proposals (support for innovation). At last, in organizations high in IOI dimensions –and above all expressing strong Team support (that is having helpful, trustworthy and cooperative colleagues)– teammates also feel a psychologically safe climate, based on non-threatening and not-judging interpersonal relationships.

Furthermore, all the IOI dimensions resulted positively related with the innovation outcome indicators. As expected, organizations highly oriented toward innovation seemed to be also more effective in accomplishing their goals and innovating. They also express a higher capability to adopt innovation, having recently at the time of the research introduced concrete product or service innovation. These results provide some evidence to the innovative orientation conceptualization as a factor that creates premises and conditions for a better performance and the effective innovation implementation.

Practical implications

Evidence from this study suggests that the IOI is a reliable and valid instrument also in its Italian version, and may therefore be adopted in researches and surveys on organizational innovation, in Italian firms as well. The IOI multidimensionality, in addition, enables to simultaneously

detect the role played by the different orientations toward innovation, at different level of analysis (organization, team, task), thus providing a dynamic and analytic perspective to understand the "state of the art" about the firm's innovativeness. Besides, following the author's suggestions (Tang, 1999), the IOI allows drawing the profile of an organization, assessing and monitoring strength and weakness areas (also benchmarking against their level in the past or those of other companies), hence raising awareness and consequently suggesting operational guidelines or intervention programs to enhance the organization's orientation to innovation.

Finally this study, consistent with previous research (Prajogo & Ahmed, 2006; Prajogo et al., 2004; Siguaw et al., 2006; Tang, 1999), confirms the orientation to innovation positive implication for the firms' performance and capability to innovate, thus offering some insights to enhance innovation. In fact, results suggest that it is useful focusing on these cultural factors, which lead to innovativeness. Accordingly, managers have to bear in mind that all the facets composing the orientation to innovation have to be nurtured, representing a strategic precursor of concrete innovation implementation.

Study limitations

We are aware of a number of limitations of our study. Indeed, caution is recommended before generalizing our findings, due to the specific cultural context where the research was conducted, the unique use of self-report data and the lack of objective outcomes. Future studies should corroborate the above findings with cross-cultural comparisons to ascertain the generalizability of findings across different cultures. Likewise, further evidence may come from multi-informant approaches and objective indicators of innovation outputs and also organizational and contextual parameters.

Although additional work is needed, particularly in the methodological domain, the results reported are promising. Indeed, this study is a first test for the IOI concurrent validation and offers a contribute to the convergent validity of the orientation toward innovation as a multidimensional construct, albeit in the future it would be worthwhile to focus on their discriminant validity, to understand whether and how some of these factors exert a specific contribute. Overall, results suggest evidence that the IOI is a reliable and valid instrument and, combined with the above recommendations, may therefore be adopted in studies on organizational innovation.

References

ADAMS, R., BESSANT, J. & PHELPS, R. (2006). Innovation management measure: A review. *International Journal of Management Review*, *8*(1), 21-47.

ALIAGA, O. (2005). *A study of innovative Human Resource Development practices in Minnesota companies*. University of Minnesota, PhD thesis.

AMABILE, T.M. (1997). Motivating creativity in organizations: On doing what you love and loving what you do. *California Management Review*, *40*(1), 39-58.

ANDERSON, N.R. & WEST, M.A. (1996). The Team Climate Inventory: The development of the TCI and its applications in teambuilding for innovativeness. *European Journal of Work and Organizational Psychology*, *5*, 53-66.

ANDERSON, N.R. & WEST, M.A. (1998). Measuring climate for work group innovation: Development and validation of the Team Climate Inventory, *Journal of Organizational Behavior*, *19*, 235-258.

ASHILL, N.J., CARRUTHERS, J. & KRISJANOUS, J. (2006). The effect of management commitment to service quality on frontline employee's affective and performance outcomes: An empirical investigation on New Zeland public healthcare sector. *International Journal of Nonprofit and Voluntary Sector Marketing*, *11*(4), 271-287.

BAIN, P., MANN, L. & PIROLA-MERLO, A. (2001). The innovation imperative. The relationships between team climate, innovation, and performance in research and development teams. *Small Group Research*, *32*(1), 55-73.

BYRNE, B.M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming. Multivariate applications series*. New York: Routledge.

BERTHON, P., HULBERT, J. & PITT, L. (1999). To serve o to create? Strategic orientations toward customer and innovation. *California Management Review*, *42*(1), 37-58.

CALANTONE, R., GARCIA, R. & DROGE, C. (2003). The effects of environmental turbulence on new product development strategy planning. *Journal of Product Innovation Management*, *20*, 90-103.

CARTER,S. & WEST, M. (1998). Reflexivity, effectiveness, and mental health in BBC-TV production teams. *Small Group Research*, *29*(5), 583-601.

CEPEDA-CARRION, G., CEGARRA-NAVARRO, J. & JIMENEZ-JIMENEZ D. (2011). The effect of absorptive capacity on innovativeness: Context and information systems capability as catalysts. *British Journal of Management*, *22*, 1-20.

CURRAL, L., FORRESTER, R., DAWSON, J. & WEST, M. (2001). It's what you do and the way that you do it: Team task, team size, and

innovation-related group processes.*European Journal of Work and Organizational Psychology*, *10*(2), 187-204.

de JONG, J. & den HARTOG, D. (2007). How leaders influence employees' innovative behaviour. *European Journal of Innovation Management*, *10*(1), 41-64.

EKVALL, G. (1996). Organizational climate for creativity and innovation. *European Journal of Work and Organizational Psychology*, *5*, 105-123.

FORNELL, C. & LARCKER, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50.

HULSHEGER, U.R., ANDERSON, N. & SALGADO, J. (2009). Team-level predictors of innovation at work: A comprehensive meta-analysis spanning three decades of research. *Journal of Applied Psychology*, *94* (5), 1128-1145.

HULT, G.T., HURLEY, R.F. & KNIGHT, G.A. (2004). Innovativeness: Its antecedents and impact on business performance. *Industrial Marketing Management*, *33*, 429-438.

HURLEY, R.F. & HULT, G.T. (1998). Innovation, market orientation, and organizational learning: An integration and empirical examination. *Journal of Marketing*, *62*, 42-54.

HURLEY, R.F., HULT, G.T. & KNIGHT, G.A. (2005). Innovativeness and capacity to innovate in a complexity of firm-level relationship: A response to Woodside (2004). *Industrial Marketing Management*, *34*, 281-283.

KONCZAK, L.J., STELLY, D.J. & TRUSTY, M.L. (2000). Defining and measuring empowering leader behaviors: Development of an upward feedback instrument. *Educational and Psychological Measurement*, *60*(2), 301-313.

LYNCH, P., WALSH, M.M. & HARRINGTON, D. (2010). *Defining and dimensionalizing organizational innovativeness*. International CHRIE Conference.

MANU, F. (1992). Innovation orientation, environment and performance: A comparison of U.S. and European markets. *Journal of International Business Studies*, *23*(2), 333-359.

MEADE, A.W., JOHNSON, E.C. & BRADDY, P.W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *The Journal of Applied Psychology*, *93*(3), 568-592.

MENGUC, B. & AUH S. (2002). Creating a firm-level dynamic capability through capitalizing on market orientation and innovativeness. *Journal of Academy Marketing Science*, *34*(1), 1552-7824.

PALEO, I.O. & WIJNBERG, N.M. (2008). Organizational output innovativeness: A theoretical exploration, illustrated by a case of popular music festival. *Creativity and Innovation Management*, *17*(1), 3-13.

PARRIS, D.L. & PEACHEY, J.W. (2013). A systematic literature review of servant leadership theory in organizational contexts. *Journal of Business Ethics*, *113*(3), 377-393.

PRAJOGO, D. & AHMED, P. (2006). Relationships between innovation stimulus, innovation capacity, and innovation performance. *R&D Management*, *36*(5), 499-515.

PRAJOGO, D., POWER, D. & SOHAL, A. (2004). Trading partner relationships on determining innovation performance: An empirical examination. *European Journal of Innovation Management*, *7*(3), 178-186.

RAYKOV T. & MARCOULIDES, G.A. (2011). *Introduction to Psychometric Theory*. New York: Routledge.

SIGUAW, J.A., SIMPSON, P.M. & ENZ, C.A. (2006). Conceptualizing innovation orientation: A framework for study and integration of innovation research. *The Journal of Product Innovation Management*, *23*, 556-574.

SIMPSON, P.M., SIGUAW, J.A. & ENZ, C.A. (2006). Innovation orientation outcomes: The good and the bad. *Journal of Business Research,*, *59*, 1133-1141.

TANG, H.K. (1998). An integrative model of innovation in organizations. *Technovation*, *18*(5), 297-309.

TANG, H.K. (1999). An inventory of organizational innovativeness. *Technovation*, *19*(1), 41-51.

VAN DE VEN, A.H., 1986. Central problems in the management of innovation. *Management Science*, *32* (5), 590-607.

VAN DIERENDONCK, D. (2011). Servant leadership: A review and synthesis. *Journal of Management*, *37*(4), 1228-1261.

WANG, C. & AHMED, P. (2004). The development and validation of the organisational innovativeness construct using confirmatory factor analysis. *European Journal of Innovation Management*, *7*(4), 303-313.

WEST, M. & ANDERSON, N.R. (1996). Innovation in top management teams. *Journal of Applied Psychology*, *81*, 680-693.

WOODSIDE, A.G. (2005). Firm orientations, innovativeness, and business performance: Advancing a system dynamics view following a comment on Hult, Hurley, and Knight's 2004 study. *Industrial Marketing Management*, *34*, 275-279.

ZALTMAN, G., DUNCAN, R. & HOLBECK J. (1973). *Innovations and Organizations*. New York (NY): John Wiley & Sons.

ZHOU, K.Z., GAO, G.Y. & YANG, Z. (2005) Developing strategic orientation in China: Antecedents and consequences of market and innovation orientations. *Journal of Business Research*, *58*, 1049-1058.

APPENDIX

IOI– Inventory of Organizational Innovativeness theoretical dimensions and Italian version

Di seguito sono elencati alcuni comportamenti che descrivono la vita di un'organizzazione. In che misura li sente corrispondenti a quanto accade nella sua organizzazione?

Item number	Dimension	English and Italian (in italic) items
1	LEA-1	Our top managers are approachable and communicative. *Il nostro management è disponibile e aperto alla comunicazione.*
2	LEA-2	Our supervisors often challenge us to be more innovative and resourceful. *I nostri responsabili ci spronano spesso a essere più innovativi e intraprendenti.*
3	LEA-3	Our top managers show great enthusiasm for innovation and work improvement. *Il nostro management mostra entusiasmo per le innovazioni e i miglioramenti sul lavoro.*
4	LEA-4 (R)	Our top managers don't value employees' opinions much. *Il nostro management non tiene molto in considerazione le opinioni dei dipendenti.*
5	SUPP-1 (**)	My organization has active programs to upgrade employees' knowledge and skills. *La mia organizzazione ha programmi concreti per l'aggiornamento delle conoscenze e delle abilità dei dipendenti.*
6	SUPP-2	There are many opportunities to exchange and generate ideas in my organization. *Nella mia organizzazione ci sono molte opportunità per scambiare e generare nuove idee.*
7	SUPP-3	My organization recognizes and rewards innovative and enterprising employees. *La mia organizzazione apprezza e premia i dipendenti innovativi e intraprendenti.*
8	SUPP-4	My organization gives adequate resources to exploring and implementing innovative ideas. *La mia organizzazione offre risorse adeguate per la ricerca e lo sviluppo di idee innovative.*
9	SUPP-5	In my organization innovative and enterprising employees are well paid. *Nella mia organizzazione i dipendenti innovativi e intraprendenti sono ben remunerati.*
10	SUPP-6 (**)	My work schedule allows me time to think of creative solutions to problems. *Il mio orario di lavoro mi consente di pensare a soluzioni creative per i problemi.*
11	SUPP-7 (**)	Innovation is clearly a part of my organization's mission or basic beliefs. *L'innovazione è chiaramente parte della mission della mia organizzazione o dei suoi principi di base.*
12	TASK-1	My work is intellectually stimulating and challenging. *Il mio lavoro è intellettualmente stimolante e sfidante.*
13	TASK-2 (**)	There are many opportunities and freedom in my work to explore and try out new ideas. *Nel mio lavoro ci sono molte opportunità e margini di libertà per cercare e sperimentare nuove idee.*
14	TASK-3	I frequently encounter non-routine and challenging work in my organization. *Nella mia organizzazione affronto spesso compiti non di routine e sfidanti.*

continued on next page

continued

Item number	Dimension	English and Italian (in italic) items
15	TASK-4 (R)	The type of work we do requires very little imagination and creativity. *Il tipo di lavoro che svolgiamo richiede davvero poca immaginazione e creatività.*
16	TASK-5	There's much knowledge to gain from the work I do for my organization. *Posso incrementare molto il mio bagaglio di conoscenza con il lavoro che faccio per la mia organizzazione.*
17	BEH-1	I found my colleagues very helpful when I encounter difficulties with my work. *Quando incontro difficoltà nel mio lavoro, i colleghi mi sono di grande aiuto.*
18	BEH-2 (R)	In my organization people show little interest in each other's work. *Nella mia organizzazione le persone mostrano poco interesse verso il lavoro dei colleghi.*
19	BEH-3	I find my colleagues very helpful in sharing knowledge and information. *Ritengo che i miei colleghi siano di grande aiuto nel condividere conoscenze e informazioni.*
20	BEH-4 (R) (**)	In my organization very few people take the initiatives to raise new projects. *Nella mia organizzazione pochissime persone prendono l'iniziativa di sviluppare nuovi progetti.*
21	INT-1 (R)	Teamwork is poor in my organization. *Nella mia organizzazione il lavoro di gruppo è scarso.*
22	INT-2	In my organization different departments work together harmoniously. *Nella mia organizzazione le diverse unità lavorano assieme in armonia.*
23	INT-3	In my organization there is a strong sense of mutual trust. *Nella mia organizzazione c'è un forte senso di fiducia reciproca.*
24	INT-4 (R)	My organization is unable to accumulate knowledge or learn and benefit from experience. *La mia organizzazione è incapace di accumulare conoscenze o di imparare e trarre beneficio dall'esperienza.*
25	RAP-1	My organization actively collects ideas for improvements from employees. *La mia organizzazione raccoglie attivamente dai dipendenti idee per il proprio sviluppo.*
26	RAP-2	In my organization employees are active in making suggestions about work improvement. *Nella mia organizzazione i dipendenti sono attivi nel proporre suggerimenti per migliorare le attività lavorative.*
27	RAP-3	In my organization there are ways to support unplanned but worthwhile initiatives. *Nella mia organizzazione ci sono dei modi per sostenere le iniziative non pianificate ma meritevoli.*
28	RAP-4	My organization evaluates project proposals with an open but pragmatic mind. *La mia organizzazione valuta le proposte di nuovi progetti con mente aperta, anche se con concretezza.*
29	RAP-5 (**)	In the pursuit of innovation or new business, my organization tolerates mistakes. *La mia organizzazione tollera gli errori, se commessi per promuovere l'innovazione o nuovi business.*
30	RAP-6	If my new idea is not accepted I can try out elsewhere in organization. *Se una mia nuova idea non è accettata, posso proporla in altri contesti dell'organizzazione.*

continued on next page

continued

Item number	Dimension	English and Italian (in italic) items
31	DOP-1	Projects and jobs are well organized and executed in my organization. *Progetti e attività sono ben organizzati e realizzati nella mia organizzazione.*
32	DOP-2	In my organization projects start with clear objectives, schedule and resource requirements. *Nella mia organizzazione i progetti partono con obiettivi e tempi chiari e con risorse adeguate.*
33	DOP-3	Projects are monitored and reviewed regularly. *I progetti sono monitorati e revisionati con regolarità.*
34	DOP-4	My organization learns about what was done right or wrong at the end of each project. *La mia organizzazione apprende dagli errori o dalle cose ben fatte, alla fine di ogni progetto.*
35	DOP-5	My organization has clearly defined achievement goals and strategic directions. *La mia organizzazione ha obiettivi e direttive strategiche chiaramente definiti.*
36	KNS-1 (**)	My colleagues and I are able to come up with creative ideas when we face tough problems. *Io e i miei colleghi siamo capaci di sviluppare idee creative per far fronte ai problemi.*
37	KNS-2 (**)	My organization creates its own intellectual assets, e.g. special techniques, patents. *La mia organizzazione sviluppa da sé le proprie risorse intellettuali (es. brevetti, tecniche particolari).*
38	KNS-3	In my organization there are many employees with strong knowledge and skills. *Nella mia organizzazione molti dipendenti hanno consistenti conoscenze e capacità.*
39	KNS-4	I have colleagues who impress me with their innovative ideas, energy, and resourcefulness. *Ho colleghi che mi colpiscono per le loro idee innovative, energia e quantità di risorse.*
40	KNS-5	I have colleagues who help others to turn ideas into action and reality. *Ho colleghi che aiutano gli altri a rendere operative e reali le loro idee.*
41	IC-1	In my organization the dissemination of information relevant to work is excellent. *Nella mia organizzazione le informazioni utili al lavoro sono divulgate in modo eccellente.*
42	IC-2	Documentation, information and databases are well managed in my organization. *Nella mia organizzazione la documentazione, le informazioni e le banche dati sono ben gestite.*
43	IC-3	My organization's information system is a great aid to finding ideas and opportunities. *Il sistema informativo della mia organizzazione è di grande aiuto per trovare nuove idee e opportunità.*
44	IC-4	My organization captures information diligently from external sources, e.g. customers. *La mia organizzazione è attenta a cogliere informazioni dall'esterno (es. dai clienti).*
45	SASS-1 (*)	My organization is effective in innovating. *La mia organizzazione è efficace nell'innovare.*
46	SASS-2 (*)	Overall, my organization is an effective organization. *Complessivamente, la mia è un'organizzazione efficace.*

Note: Items refer to the following theoretical dimensions: LEA = leadership; SUPP = support; TASK = task; BEH = behaviour; INT = integration; RAP = raising projects; DOP = doing projects; KNS = knowledge and skills; IC = information and communication; SASS = summary assessment items about general perception of organizational innovativeness and effectiveness.
(R) Reversed items.
(*) Items excluded from the factorial analysis because not specific to any scales.
(**) Deleted items.