ISSN 1927-0526 (Print)
ISSN 1927-0534 (Online)

Journal of Educational and Developmental Psychology

Vol. 7, No. 2 November 2017

CANADIAN CENTER OF SCIENCE AND EDUCATION

Editorial Board

Contents

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

The Fibonacci Life Chart Method as a Predictor of Spiritual Experience

Robert G. Sacco[1]

[1] Toronto, Canada

Correspondence: Robert G. Sacco, Toronto, Canada. E-mail: robgsacco@gmail.com

Received: March 15, 2017 Accepted: April 5, 2017 Online Published: April 12, 2017

doi:10.5539/jedp.v7n2p1 URL: http://doi.org/10.5539/jedp.v7n2p1

Abstract

The Fibonacci Life Chart Method (FLCM) provides a framework linking development and spirituality. This study addressed the need for empirical research to test the hypotheses proposed by Sacco (Sacco, 2016). To address this problem, this study used case reports (N = 196) from the Alister Hardy Religious Experience Research Centre. The dynamical aspects of ages 11, 18, and 30 were examined as predictors of increased spiritual experience in adolescents and young adults. Results showed only ages 17 and 18 predicted a higher frequency of spiritual experience between ages 11 and 35. Age 18 was associated with a higher effect size (r = .27). This finding provides some empirical support for the FLCM as a predictor of spiritual experience, but not all hypotheses found support. Limitations to the study's design are discussed along with implications for future research.

Keywords: dynamical systems theory, Fibonacci numbers, Fibonacci Life Chart Method, golden ratio, spiritual experience, synchronicity

1. Introduction

Spirituality and religion are central to people worldwide. In the United States, polls show 89% of respondents believe in God and most are religious, spiritual, or both (Fuller, 2001; Pew Research Center, 2014). Spiritual and religious experiences are also common in the general population. Studies show about 35% of people have had a spiritual or religious experience (Greeley, 1975; Hay, 1990). The widespread occurrence of spiritual experience contrasts with a lack of research into these complex human experiences (Hood & Zhuo, 2013). The lack of research, particularly among younger groups, is problematic. Researchers have shown that emerging adults experience a high number of spiritual struggles and are at risk for negative psychological outcomes (Bryant & Astin, 2008; Exline & Rose, 2013). This raises questions about whether some spiritual experiences and struggles might be part of a normal developmental process among young adults. Understanding predictors, correlates, and mechanisms of spiritual experience may thus help to improve spiritual struggles.

Adolescence and emerging adulthood provide a useful context to examine the dimensions of spiritual experience, for two reasons. First, studies show that adolescence and emerging adulthood are important developmental periods for spiritual experience. For instance, Smith (2006) found the average age of spiritual transformation was 30 years and signified that early adulthood is the most common period. Indeed, 60% of subjects reported they were under 29 when their spiritual transformation occurred (Smith, 2006). Second, spiritual experience appears linked to identity growth during late adolescence and emerging adulthood. As Erikson proposed personality develops in eight epigenetic stages over the course of the life span (Erikson, 1982; Sacco, 2013). The main developmental task of late adolescence and emerging adulthood is to establish a personal identity. This suggests that developmental changes in identity might explain changes in spirituality. Thus, the developmental origins of spiritual experience deserve further attention.

To better understand the relationship between age and spiritual experience, Sacco (2016) described a new method, termed the Fibonacci Life Chart Method (FLCM). Derived from dynamical systems theory, the FLCM provides a mathematical framework for human development (Sacco, 2013). The FLCM outlines eight stages in human development based on the Fibonacci sequence. Waskom (1972) first theorized that human development stages might follow the Fibonacci sequence. To describe human development with the Fibonacci sequence, Waskom asserted the numbers that mark human developmental stages are the same numbers expressed in the

Fibonacci sequence: 1, 5, 8, 13, 21, 34, and 55 (Rose, 1991). In contrast, Sacco (2013) suggested it was not clear why Fibonacci numbers must have numeric values in years. Sacco claimed a more useful taxonomy of the Fibonacci numbers is the 24-hour day/night cycle. For example, human embryonic cells have a doubling time of 24 hours (Lagarkova, Eremeev, Svetlakov, Rubtsov, & Kiselev, 2010). Yet whether the FLCM based on a 24-hour cycle has a direct effect on human outcomes has not been shown.

The importance of the Fibonacci numbers comes from several sources. Many physical and biological systems have structures that approximate Phi, an irrational number (about 1.618034), also known as the "golden ratio". The relation between Fibonacci numbers and the golden ratio is that the ratios of the successive numbers in the Fibonacci sequence converge on the golden ratio (Grattan-Guinness, 2002). Harmonic characteristics related to the golden ratio appear in the arrangement of leaves on a plant stem (Okabe, 2011), spiral structures of galaxies (Grattan-Guinness, 2002), pulse frequency of a star (Lindner et al., 2015), quantum phase transition (Coldea et al., 2010), nucleotide frequencies (Yamagishi & Shimabukuro, 2008), and cell (Staff et al., 2012) and shell (Gosling, 2008) growth. In human science, the golden ratio has been observed in body proportions (Ferring & Pancherz, 2008), bronchial airway segment bifurcations (Goldenberger, West, Dresselhaus, & Bhargava, 1985), gait phases of walking (Iosa et al., 2013), hair whorls (Paul, 2016), and aesthetic preference (Ricketts, 1982; Russell, 2000).

1.1 The FLCM and Spiritual Experience

Of possible further significance of the golden ratio to human perception might be spiritual and religious experience. As William James (1902/1985) noted over 100 years ago, spiritual experiences are diverse. Most modern researchers would agree that religion and spirituality overlap, but are distinct constructs (e.g., Piedmont, 2005; Zinnbauer & Pargament, 2005). Religiosity involves participation in collective beliefs, activity, and ceremonies of traditional organized religions. A *religious experience* therefore relates to a particular religious doctrine (e.g., listening to a sermon, reading the Bible, or alone in church). Spirituality involves a personal subjective experience of the sacred. A *spiritual experience* therefore transcends the ordinary aspects of daily life. More common spiritual experiences encompass awe and joy that lifts one out of the mundane (Underwood & Teresi, 2002). Spiritual experiences, however, also relate to more paranormal phenomena (e.g., Hood & Zhuo, 2013; Kelly & Tucker, 2015). A *paranormal experience* deviates from "every day" causal principles.

Hardy's (1979) analysis of 3,000 case reports classified spiritual or religious experience into eight major types: *synchronicity and patterning of events*, *presence of God*, *answered prayer*, *presence not called God*, *sacred presence in nature*, *experiencing that "all things are one"*, *presence of the dead*, and *presence of evil*. Furthermore, "synchronicity and patterning of events" is the most commonly reported category of spiritual experience (Hay, 1990). Of the people who reported experience of synchronicity, only 32% regarded this as religious (Hay, 1990). In addition to Hardy's categories, two more types of classification based on the context of the experience are: (a) crisis experiences, and (b) non-crisis experiences. Distinct differences exist between crisis experiences and non-crisis experiences. For example, a crisis can often stimulate creative thinking or encourage new goals. Hence, crisis experiences may emerge more "spontaneously" resulting not from specific practices, but as resolution and solutions to problems (Batson, Schoenrade, & Ventis, 1993). Also, a crisis often involves emotions of fear, grief, or other powerful emotions.

The claim the golden ratio might relate to spiritual experience is made more plausible by three sets of empirical findings. First, spiritual experiences tend to occur in contexts of stress and crisis (James, 1902/1985; Smith, 2006). From a dynamical systems perspective, stress or crisis can act as a perturbation causing a change in system parameters. Complex systems resist changes to their attractor states. However, when the perturbation is strong enough, a system can move from a more stable dynamical attractor pattern to a less stable attractor pattern (Thelen & Smith, 1994). Such a transition in a complex system is referred to as a "phase transition" a brief chaotic period. Phase transitions are critical points in dynamic systems. They also relate to the Fibonacci sequence (Linage, Montoyaa, Sarmientob, Showalter, & Parmananda, 2006). Specifically, the Fibonacci sequence appears within the Feigenbaum scaling of the period doubling cascade to chaos (Linage et al., 2006). Thus, the Fibonacci sequence is at the root of dynamical phase transition.

Second, at the core of spiritual experience is a sense of unity (Hardy, 1979; James, 1902/1985). Why might spiritual experience relate to a sense of unity? It may be helpful to consider the dynamical aspects of phase transitions. Empirically, the critical dynamics of phase transitions are shown to have nonlocal and fractal connectivity. Nonlocality refers to correlations between spatially separated events (Stapp, 2009). A fractal is a symmetry having a pattern that repeats at different scales (Bak, 1996). Significantly, fractal geometry includes

the Fibonacci sequence as a unifying theme (Devaney, 1999). Perhaps, too, there is a unity found in nonlocality, fractals, the Fibonacci sequence, and spiritual experience (Sacco, 2016). If spiritual experiences involve emergent dynamics within a stressed state (e.g., Smith, 2006), then it is logical to assume that spiritual experience relates to nonlocality and fractals generally, and particularly experiences of ultimate meaning, unity, and interconnectedness (Hogenson, 2005, 2014; Sacco, 2016).

Third, people often describe the feeling of spiritual experience as beauty, harmony, and love. Likewise, aesthetic consideration has surrounded the golden ratio ever since humans first reflected on the geometric forms of the world. The golden ratio appears in the sacred art of Egypt, India, China, Islam and other traditional civilizations (Livio, 2008). The golden ratio was called the "divine proportion" during the Renaissance, from the title of a book published by the mathematician Luca Pacioli in 1509 (Huntley, 1970), and illustrated by Leonardo da Vinci. The golden ratio, or divine proportion, has for centuries represented perfect harmony, or the most attractive proportion in almost all things. Thus, the same sense of harmony perceived in the dimension of space may also be perceived in time (Sacco, 2016). As shown above, there are reasonable theoretical grounds that FLCM could provide a model for researchers to explore spiritual experience.

1.2 The Present Study

The goal of this study was to examine how dynamical aspects of FLCM ages 11, 18, and 30 may predict increased spiritual experiences. This study focused on adolescence and emerging adulthood as important developmental periods for spiritual experience. The study data is from the Alister Hardy Religious Experience Research Centre (RERC), an extensive database of over 6,000 accounts of spiritual experience collected since 1969. To examine the adolescence and emerging adulthood period only data between ages 11 and 35 was used. Based on the FLCM it was hypothesized ages 11, 18, and 30 may predict a higher frequency of spiritual experience.

2. Method

2.1 Study Samples

The Study Sample comprised accounts of religious or spiritual experience received by the Alister Hardy Religious Experience Research Centre (RERC) in response to mass-media appeals. Respondents were from different regions around the world and religious backgrounds. The entire database comprised 6,482 respondents. With access to such a vast and diverse pool of subjects, it was possible to select test cases strategically. The inclusion/exclusion criteria were:

(1) Single experience: Respondents with multiple experiences were excluded as single experiences would mean the most memorable experience.

(2) Adequate age specification: Some respondents provided an age range (e.g., "21-23") or life stage (e.g., "teens, young adult"), and not a single discrete age. This data was excluded as it did not enable measures of central tendency.

(3) Adolescence and emerging adulthood: This paper defined adolescence and emerging adulthood as persons between ages 11 and 35. Therefore, age of writing and age of experience was restricted between these ages.

The database search spanned all years since 1969 up to the present study. The electronic search yielded 196 case reports that met inclusion criteria and thus warranted further review. Table 1 shows application of the exclusion criteria. Collection of demographic data also allowed an analysis of the representativeness of the Sample.

Table 1. Exclusion criteria for the study of spiritual experience

	Available for analysis	Excluded *n*	Included *n*
	N = 6,482		
Age of experience			
Not specified		3,024	
Multiple ages/not adequately specified		1,888	
Age of experience < 11		78	
Age of experience > 35		551	

Age of writing			
	Not specified	103	
	Not adequately specified	50	
	Age of writing < 11	1	
	Age of writing > 35	591	
			196

2.2 Materials

The Fibonacci Life Chart Method (Sacco, 2013, 2016) was used to generate chronological ages. Table 2 shows calculation of FLCM chronological ages. As shown in Table 2, the FLCM generates a pattern characterized by eight age groups: early infancy (0-2), toddler (2-4), early childhood (4-7), middle childhood (7-11), adolescence (11-18), young adulthood (18-29), middle adulthood (29-48), and older adulthood (48-78+) (Sacco, 2013). The model hypothesizes that these ages are biological phase transitions and marked by potential changes in identity.

Table 2. The Fibonacci Life Chart Method

Fn	Date	Chronological Age
0	January 1, 2000	0.00
1	January 2, 2000	0.00
1	January 3, 2000	0.01
2	January 5, 2000	0.01
3	January 8, 2000	0.02
5	January 13, 2000	0.03
8	January 21, 2000	0.05
13	February 3, 2000	0.09
21	February 24, 2000	0.15
34	March 29, 2000	0.24
55	May 23, 2000	0.39
89	August 20, 2000	0.64
144	January 11, 2001	1.03
233	September 1, 2001	1.67
377	September 13, 2002	2.70
610	May 15, 2004	4.37
987	January 27, 2007	7.08
1597	June 12, 2011	11.45
2584	July 9, 2018	18.53
4181	December 19, 2029	29.99
6765	June 27, 2048	48.52
10946	June 16, 2078	78.51

Note. Fibonacci numbers represent 24-hour days.

2.3 Search Procedure

The Alister Hardy electronic database was comprehensively searched. The data reported in this paper was collected from July 2016 to September 2016. No date restrictions or keywords were imposed on the searches.

Case reports were generated only with the main search box button, which retrieved the entire database of 6,482 case reports. An automated web scraping system was used to extract data from the webpages (Landers, Brusso, Cavanaugh, & Collmus, 2016). The extracted data was exported to Microsoft Excel (version 2013; Microsoft, Redmond, WA). With the scraped dataset, sorting of age variables was performed with the filter function in Microsoft Excel.

Case reports were filtered based on: (a) single experiences, (b) adequate age specification, and (d) age of experience and age of writing ≥ 11 and ≤ 35. Ages could be rapidly sorted as single integers with the filter function in Microsoft Excel. This sorted all missing ages, or those not adequately specified, into a separate column. For example, to the age of experience question, some respondents provided no answer. Others answered with multiple ages (e.g., "12, 22, 31"), an age range (e.g., "21-23"), or life stage (e.g., "teens, young adult"). Others were unsure (e.g., "about 20", "11 or 12"). All irrelevant case reports were excluded from further analysis. This created a column of single age integers for analysis and comparison. The exclusion criteria resulted in 196 relevant case reports (Table 1). All ages were then sorted from low numbers to high numbers. Ages were then compared for age matches. A match corresponded to the age of experience equaling age 11, 18, or 30. The Microsoft Excel auto-sum function totaled up age matches.

2.4 Statistical Analysis

Statistical analysis of the data was performed using Microsoft Excel (version 2013; Microsoft, Redmond, WA) and GraphPad Prism version 7.00 for Windows (GraphPad Software, San Diego, California, USA). Microsoft Excel generated descriptive statistics, while GraphPad Prism produced inferential statistics. The chi-square test for goodness of fit was used to compare observed and expected frequencies. The chi-square test can satisfactorily compare the observed distribution of data to an expected distribution. Statistical significance was set at $p < .05$.

3. Results

Of the 196 case reports, demographic results show that most respondents were female (62%), Christian (68%), and were from the United Kingdom (65%) (Table 3). The age of writing range was 14 to 35 ($M_{age\ of\ writing} = 27.13$). The age of experience range was 11 to 35 ($M_{age\ of\ experience} = 21.46$). Mean temporal distance between age of writing and age of experience was 5.67 years. The probability distribution for age of experience is shown in Figure 1. As shown in Figure 1, ages 17 and 18 show a significant peak. Age 18 has the highest reporting level compared to the other ages.

Table 3. Demographic characteristics of the study sample

Variable		#	%
Gender			
	Male	70	35.7%
	Female	121	61.7%
	Not specified	5	2.6%
Religious background			
	Agnostic	1	0.5%
	Atheist	3	1.5%
	Buddhist	2	1.0%
	Christian	134	68.4%
	Jewish	6	3.1%
	No Allegiance	10	5.1%
	Not specified	40	20.4%
Country			
	Other	17	8.7%
	United Kingdom	127	64.8%
	USA	29	14.8%

	Not specified	23	11.7%
Age of writing			
	< 11	--[a]	--
	11 – 19	27	14.2%
	20 – 29	86	45.3%
	30 – 35	83	43.7%
	> 35	--[b]	--
	Not specified	0	0.0%
	Mean	27.13	

Note. Other countries include: Australia, Austria, Canada, Japan, South Africa, Spain, Sweden, and Turkey.

[a] No case reports for age of writing < age 11 were included.

[b] No case reports for age of writing > age 35 were included.

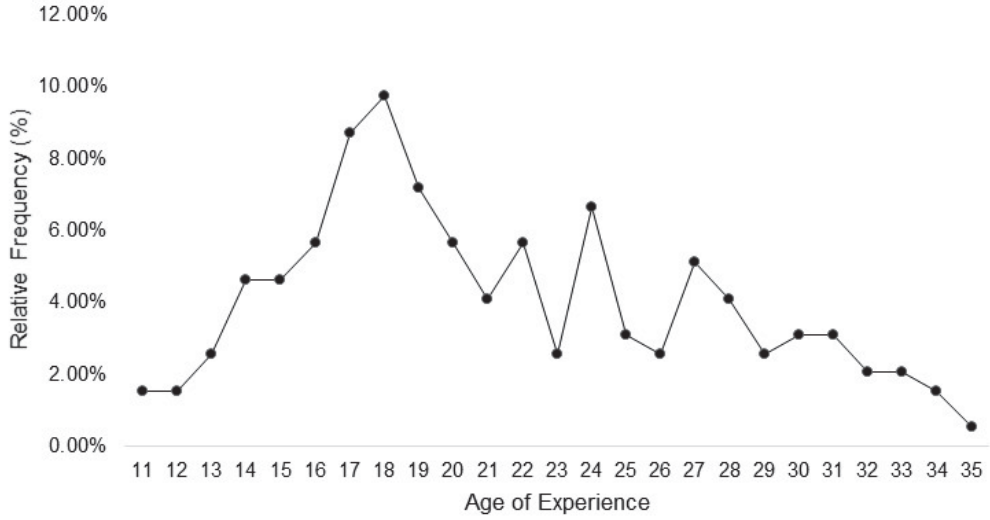


Figure 1. Distribution of spiritual experience (relative probability) as a function of reported age of experience summarized for ages 11 to 35

It was hypothesized that ages 11, 18, and 30 would predict increased spiritual experiences. To examine the association between ages 11, 18, 30 and spiritual experiences, case reports (N = 196; $M_{age\ of\ writing}$ = 27.13) were checked for age matches. Of the 196 case reports, there were: (a) 3 matches at age 11; (b) 19 matches at age 18; and (b) 6 matches at age 30. For purposes of answering the research question, the expected and observed frequencies of matches were compared (see Table 4). The expected proportion is the probability of a match occurring randomly; obtained by dividing the number of possible matches by the age range. For example, the probability for age 11 was calculated as [1 = (11)] divided by the age range [24 = (35-11)]. The expected proportion is, therefore, .042 match and .958 mismatch. Table 4 shows the chi-square contingency table analysis.

Table 4. Comparison of observed and expected distributions of spiritual experience matches

	Observed *n*	Expected *n*	Expected %
		Age 11	
Match	3	8.23	.042
Mismatch	193	187.77	.958
		Age 18	
Match	19	8.23	.042
Mismatch	177	187.77	.958
		Age 30	
Match	6	8.23	.042
Mismatch	190	187.77	.958

The results of the chi-square test for goodness of fit showed the observed frequency of matches for age 18 significantly departed from the expected distribution in a positive direction, · (N = 196) = 14.70, p < .001. The overall weighted mean effect size was .27, a medium effect size. Thus, the hypothesis that age 18 predicts an increase in the frequency of spiritual experience appears to receive strong support (p < .05). The results of the chi-square test for goodness of fit showed the observed frequency of matches for FLCM ages 11 and 30 did not significantly depart from the expected distribution, · (N = 196) = 3.47, p = .063 and · (N = 196) = 0.63, p = .427. The only other age with a significant test statistic was age 17, · (N = 196) = 9.75, p < .002. Age 17 corresponded to a smaller resulting effect size (.22). Thus, overall, the hypothesis that FLCM ages predict an increase in the frequency of spiritual experience appears to receive support only for age 18 (p < .05).

4. Discussion

In a previous paper (Sacco, 2016), it was shown how FLCM could be used as a theoretical framework for systematic empirical research into spiritual experience. In this article, the goal was to test whether the dynamical aspects of ages 11, 18, and 30 predict a higher frequency of spiritual experience. Data from this study indicate only ages 17 and 18 predict an increase in spiritual experiences among adolescents and young adults. Age 18 was associated with a higher overall effect size (r = .27). This finding supports the hypothesis that age 18 is a dynamical system attractor for spiritual experience among adolescents and young adults.

4.1 Phase Transition and Spiritual Experience

The findings presented here have implications for several broader issues. The findings suggest that personality changes between ages 17-18 may predispose some people to higher levels of spiritual experience. Adolescence and early adulthood are a unique period because of the biological maturation of different brain regions and circuits (e.g., Dumontheil, 2014). Therefore, one would expect some stability in the onset and offset of sensitive periods. This feature of sensitive periods is an essential premise of Erikson's life stage theory (Erikson, 1982). FLCM builds on Erikson's life stage theory by outlining eight stages in the human lifespan based on the Fibonacci sequence (Sacco, 2013). A basic premise of FLCM is these eight stages occur as biologically determined phase transitions from one attractor state to another. The transition marks the end of one period of stability and emergence of another (Sacco, 2013). This hypothesis is in accord with several studies that show FLCM ages (i.e., 1, 2, 4, 7, 11, 18, 30, 48, and 78) relate to significant biological and psychological changes.

For instance: (a) all major theories describe intellectual development with transitions at 1-2, 6-7, and 11-12 years associated with increasingly abstract representations (e.g., Piaget, 1970; Case, 1985), and cortical gray matter volume, reflecting neuronal density and the number of connections between neurons, peaks around age 11 (e.g., Giedd, 2004); (b) puberty begins around ages 10-11 in girls and ages 11-12 in boys (Kail & Cavanaugh, 2010) and male testosterone peaks at age 18-19 (e.g., Kelsey, Li, Mitchell, Whelan, Anderson, & Wallace, 2014); (c) the final stage of third molar development occurs at age 18 (Roberts, Lucas, Andiappan, & McDonald, 2017); (d) personality stability becomes fixed at age 30 (e.g., Terracciano, Costa, & McCrae, 2006); (e) happiness levels are lowest at age 48.5 around the world (Blanchflower & Oswald, 2008); (f) the average age at menopause marking the end of the fertile lifespan has remained constant at 50 years old (e.g., Tehrani, Solaymani-Dodaran, Tohidi,

Gohari, & Azizi, 2013); and (g) the average age of onset for dementia in males is 78.8 (Brinks, Landwehr, & Waldeyer, 2013).

Collectively, these findings imply a unique role of FLCM ages in predicting biological and psychological phase transition. These phase transitions can result from self-organization processes at multiple levels (e.g., molecules, genes, cell, organ, organ system, organism, behavior, and environment) that can influence each other (Thelen & Smith, 1994). Similarly, the association of age 18 with increased spiritual experience imply a critical period for the development of identity due to major physical, cognitive, and social changes. Theoretically, FLCM ages may be conceived as an internal timing mechanism that start a transformation of an attractor's objective phase space. Dynamical systems attractors are characterized by a complex and fractally arranged folding of space called chaotic or strange attractors (Thelen & Smith, 1994). With strange attractors, the attractor ends up filling a definite fraction of space with points (i.e., it is fractal). Thus, age 18 might be a strange attractor for spiritual experience.

4.2 The Nature of Spiritual and Religious Experience

A broader issue concerns the nature of spiritual experience. The classic text in the psychology of religion is *The Varieties of Religious Experience* (James, 1902/1985), and is subtitled "a study in human nature". This implies that psychologists ought to be interested in spiritual and religious experiences insofar as they are part of human nature. Surprisingly, theories and models of spiritual experience are largely devoid of reference to the natural sciences or human development. Dynamical systems theory is an approach that has transformed the natural sciences (Abraham & Shaw, 1992; Haken, 2002; Thelen & Smith, 1994). This perspective resonates especially well with enduring issues in spiritual experience research, and thus serves as a valuable integrative framework for research and theory construction. The dynamical perspective provides a set of principles, methods, and tools that impose precision on topics often opaque when viewed through traditional lenses.

Since the work of James (1902/1985) more than 100 years ago, psychologists have recognized the central role of stress and crisis to spiritual experience. In the current study, upon review of case reports at age 18, a large number (over 50%) of descriptions related to problems or a crisis event. Dynamical systems theory provides a scientific model for how humans self-organize and change over time (Thelen & Smith, 1994). Modern dynamical systems theory can help in understanding how spiritual experiences evolve during times of stress and crisis. For example, (a) an extreme amplification of chaos may create a crisis, (b) the process of destabilization and restabilization through self-organization, (c) the ability of self-organizing systems to "solve" problems, (d) feeling sudden change, (e) the sense of interconnectedness with the rest of reality, (f) awareness of an emergent pattern of life, (g) and perception of unity.

Clearly, the components of spiritual experience align empirically and conceptually with dynamical systems theory. Therefore, integration with dynamical systems theory can proceed from the premise that the same patterns that underlie dynamical phase transition also underlie spiritual experience (Sacco, 2016). Most importantly, FLCM predicts many spiritual experiences involve the dynamical aspects of fractals and nonlocality (Sacco, 2016). Researchers have begun to explore how synchronicity experiences could relate to quantum nonlocality (e.g., Carminati & Martin, 2008; Martin, Carminati, & Carminati, 2010). The possibility of a macroscopic cognitive effect responding to quantum effects is of interest given the new field of "quantum biology". Among the variety of biological phenomena that fall within the emerging field of quantum biology are olfaction (Turin, 2002), photosynthesis (Sarovar, Ishizaki, Fleming, & Whaley, 2010), DNA mutation (Löwdin, 1966), and the navigation of some bird species via magnetic field sensing and quantum entanglement (Rodgers & Hore, 2009).

4.3 Strengths and Limitations

This project represents one of the first attempts to predict spiritual experience developmentally. It used a comprehensive international database. Data collection included automated and targeted extraction for statistical analyses. Consequently, the overall effect of coding errors is expected to be zero. The study also benefited from a large sample size to ensure enough statistical precision. Although the present study contributes to the literature by testing the predictions of FLCM, this study is preliminary due to several limitations.

The most general limitation is the accuracy of the data. Retrospective memory is a major issue in this study. Since the respondents reflected on their personal experiences, the memories could have been inaccurate. An ongoing debate relates to the reliability of adult memory. Some research points toward the view that memory

decays with age and so memories can become distorted, but other research suggests memories are remembered well into adulthood (Kvavilashvili, Kornbrot, Mash, Cockburn, & Milne, 2009).

Also, the present sample, although international, mainly represented people from the United Kingdom or USA (79.6% of international respondents) who were Christian. Obviously, this demographic constraint limits the generalizability to other samples (e.g., non-Western, non-Christian). Future work should examine the predictive relationships between development and spiritual experience in a diverse set of religious affiliations to determine whether development impacts spiritual experience differently in different faith traditions (or between religious and non-religious people).

A further issue is construct validity. The aggregation of religious and spiritual experiences raises the concern of construct overlap, because these are considered distinct constructs (Piedmont, 2005; Zinnbauer & Pargament, 2005). This may cause misclassification that affects the representativeness of the system. In this regard, what is important is the upward or downward trends observed with the stability of this form of reporting.

Finally, the present study excluded respondents who had multiple experiences. Conventionally, when researchers ask participants to report on their personal experiences, they often ask to select one particular episode (i.e., the one they remembered best). However, categorizing data based on single experiences may not have distinguished the most important experiences. The estimated 50% trimming of data potentially affects the external validity of the study. Yet it is not expected to affect the results in a way that would alter the main conclusions of the article.

4.4 Recommendations for Future Research

In clinical settings, this work helps to understand the etiology of spiritual experience. It is important to consider the implications of spiritual experience due to the long term impact on well-being (Bryant & Astin, 2008; Exline & Rose, 2013). Evidence suggests crisis and disequilibrium are precursors to spiritual struggle, and that spiritual struggle is a precursor to spiritual growth (Desai & Pargament, 2015; Werdel, Dy-Liacco, Ciarrocchi, Wicks, & Breslford, 2014). However, if a person is not supported through this process, then the possibility for healthy growth may be hindered. This suggests early intervention is needed to prevent maladaptive behavior patterns which may accompany unresolved spiritual struggles (Dworsky, Pargament, Wong, & Exline, 2016).

To the extent that FLCM is a valid predictor of spiritual experience requires clarification. The present study provided no direct evidence of an association between ages 11 and 30 and increased spiritual experience as predicted by FLCM. However, this does not mean the ages are without importance. Instead, alternative interpretations of the results should be considered. One explanation is that certain types of spiritual experience may depend on adult-like cognitive skills such as abstract thinking (Fowler, 1995). For instance, synchronicity arises from integrating symbols and experience (Jung, 1952). This depends on abstract thinking only emerging around age 11 (e.g., Dumontheil, 2014). Thus, maturity could influence a capacity for spiritual awareness.

Another possible explanation why ages 11 and 30 did not predict increased spiritual experience includes the identity formation process. Erikson coined the term *identity crisis* to describe the uncertainty and anxiety adolescents experience as they struggle with alternatives and choices (Erikson, 1982). Erikson contended the difficult aspects of identity crises are often resolved between ages 15-18. If identity crisis and resolution predict spiritual experience, then most spiritual experiences should occur during the identity formation process. This understanding appears to fit the present data: Ages 17 and 18 had the highest reporting level compared to other ages. Presumably, because these ages represent significant developmental changes in identity.

The identity formation process may explain why ages 11 and 30 were not strong predictors of spiritual experience. Children aged 11 have neither thought much about nor resolved identity questions. While adults aged 30 are more likely to have resolved identity issues and achieved a stable identity by making personal commitments to various life domains (e.g., Terracciano, Costa, & McCrae, 2006). This stability means that after age 30, while also open and flexible, the system dynamics may crystalize in attractor states (Thelen & Smith, 1994). To support this understanding, evidence exists that personality and self-concept predict the experience of spiritual struggles (Grubbs, Wilt, Stauner, Exline, & Pargament, 2016).

To further this program of research, future studies should replicate it with various changes. One change should be to include questionnaires that address the concept of synchronicity. Researchers could administer synchronicity questionnaires to various selected populations. Prime candidates include people interested in parapsychology and spiritual groups. Specifically, it would be worthwhile to conduct this research with a group of clients involved in Jungian-oriented psychotherapy. Many examples of Jung's theories (e.g., archetypes and the collective unconscious) can support synchronicity observations. Researchers should also have subjects keep

detailed diaries of their synchronicity experiences over extended periods. This will provide an accurate source of memories for testing. These changes will increase internal and external validity.

Finally, the extent these findings generalize to FLCM secondary and tertiary date intervals requires clarification (Sacco, 2016). The dynamics of secondary and tertiary date intervals may depend on the degree of coupled interactions within a synchronized system (Hogenson, 2014). Considering these dynamics can manifest in subtle ways requires careful attention to the complex interactions among the person's development. Researchers might expect the FLCM to be more readily studied during the identity formation process between ages 18 and 29. Since interactions between development and spiritual experience may be most intense during this period.

5. Conclusion

This study examined whether the dynamical aspects of ages 11, 18, and 30 predict increased spiritual experience in adolescents and young adults. The results showed age 18 predicted increased spiritual experience. However, ages 11 and 30 did not predict increased spiritual experience. The data from this study suggest age 18 is a dynamical system attractor for spiritual experience, as predicted by the FLCM. The implication is that developmental changes in identity may be predictors of increased spiritual experience. These findings point to the intricate relationship between development and spiritual experience, and suggest the need to understand development as a contributing factor to spiritual experience.

Acknowledgments

I am grateful to the Alister Hardy Trust and the Alister Hardy Religious Experience Research Centre, University of Wales Trinity Saint David, Lampeter, UK, for permission to access their archive of spiritual accounts and publish this research. Request for copies of the study data set should be addressed to robgsacco@gmail.com.

References

Abraham, R. H., & Shaw, C. D. (1992). *Dynamics: The geometry of behavior*. Redwood City: Addison-Wesley.

Bak, P. (1996). *How nature works: The science of self-organized criticality*. New York, NY: Springer. https://doi.org/10.1007/978-1-4757-5426-1

Batson, C. D., Schoenrade, P., & Ventis, W. L. (1993). *Religion and the individual: A social psychological perspective*. New York, NY: Oxford University Press.

Blanchflower, D. G., & Oswald, A. J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine, 66*, 1733-1749. https://doi.org/10.1016/j.socscimed.2008.01.030

Brinks, R., Landwehr, S., & Waldeyer, R. (2013). Age of onset in chronic diseases: New method and application to dementia in Germany. *Population Health Metrics*, *11*(1), 6. https://doi.org/10.1186/1478-7954-11-6

Bryant, A. N., & Astin, H. S. (2008). The correlates of spiritual struggle during the college years. *The Journal of Higher Education*, *79*, 1-27. https://doi.org/10.1353/jhe.2008.0000

Carminati, G. G., & Martin, F. (2008). Quantum mechanics and the psyche. *Physics of Particles and Nuclei*, *39*(4), 560-577. https://doi.org/10.1134/S1063779608040047

Case, R. (1985). *Intellectual development: Birth to adulthood*. New York, NY: Academic Press.

Coldea, R., Tennant, D. A., Wheeler, E. M., Wawrzynska, E., Prabhakaran, D., Telling, M., ... Kiefer, K. (2010). Quantum criticality in an Ising chain: Experimental evidence for emergent E8 symmetry. *Science*, *327*(5962), 177-180. https://doi.org/10.1126/science.1180085

Desai, K. M., & Pargament, K. I. (2015). Predictors of growth and decline following spiritual struggles. *The International Journal for the Psychology of Religion*, *25*(1), 42-56. https://doi.org/10.1080/10508619.2013.847697

Devaney, R. L. (1999). The Mandelbrot set, the Farey tree, and the Fibonacci sequence. *The American Mathematical Monthly*, *106*(4), 289-302. https://doi.org/10.2307/2589552

Dumontheil, I. (2014). Development of abstract thinking during childhood and adolescence: The role of rostrolateral prefrontal cortex. *Developmental Cognitive Neuroscience*, *10*, 57-76. https://doi.org/10.1016/j.dcn.2014.07.009

Dworsky, C. K. O., Pargament, K. I., Wong, S., & Exline, J. J. (2016). Suppressing spiritual struggles: The role of experiential avoidance in mental health. *Journal of Contextual Behavioral Science*, *5*(4), 258-265. https://doi.org/10.1016/j.jcbs.2016.10.002

Erikson, E. H. (1982). *The life cycle completed*. New York, NY: Norton.

Exline, J. J., & Rose, E. D. (2013). Religious and spiritual struggles. In R. F. Paloutzian, & C. Park (Eds.), *Handbook of the psychology of religion and spirituality* (2nd ed., pp. 380-398). New York, NY: Guilford. https://doi.org/10.1037/14045-025

Ferring, V., & Pancherz, H. (2008). Divine proportions in the growing face. *American Journal of Orthodontics and Dentofacial Orthopedics*, *134*(4), 472-479. https://doi.org/10.1016/j.ajodo.2007.03.027

Fowler, J. W. (1995). *Stages of faith: The psychology of human development and the quest for meaning*. New York, NY: Harper One.

Fuller, R. C. (2001). *Spiritual, but not religious: Understanding unchurched America*. New York, NY: Oxford University Press. https://doi.org/10.1093/0195146808.001.0001

Giedd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Annals of the New York Academy of Sciences*, *1021*(1), 77-85. https://doi.org/10.1196/annals.1308.009

Goldenberger, A. L., West, B. J., Dresselhaus, T., & Bhargava, V. (1985). Bronchial asymmetry and Fibonacci scaling. *Cellular and Molecular Life Sciences*, *41*, 1537-1538. http://dx.doi.org/10.1007/BF01964794

Gosling, E. (2008). *Bivalve molluscs: Biology, ecology and culture*. Oxford, UK: Blackwell Science.

GraphPad Prism [Computer software]. (2016). Retrieved from http://www.graphpad.com

Grattan-Guinness, I. (Ed.). (2002). *Companion encyclopedia of the history and philosophy of the mathematical sciences*. London, UK: Routledge.

Greeley, A. (1975). *The sociology of the paranormal*. Beverly Hills, California: Sage Publications.

Grubbs, J. B., Wilt, J., Stauner, N., Exline, J. J., & Pargament, K. I. (2016). Self, struggle, and soul: Linking personality, self-concept, and religious/spiritual struggle. *Personality and Individual Differences*, *101*, 144-152. https://doi.org/10.1016/j.paid.2016.05.365

Haken, H. (2002). *Brain dynamics*. Berlin, Heidelberg: Springer.

Hardy, A. (1979). *The spiritual nature of man: A study of contemporary religious experiences*. New York, NY: Oxford University Press.

Hay, D. (1990). *Religious experience today: Studying the facts*. London, UK: Mowbray.

Hogenson, G. B. (2005). The self, the symbolic and synchronicity: Virtual realities and the emergence of the psyche. *Journal of Analytical Psychology*, *50*(3), 271-284. https://doi.org/10.1111/j.0021-8774.2005.00531.x

Hogenson, G. B. (2014). Are synchronicities really dragon kings? In H. Atmanspacher, & C. A. Fuchs (Eds.), *The Jung-Pauli Conjecture and its Impact Today* (pp. 201-215). Exeter, UK: Imprint Academic.

Hood, R. W., & Zhuo, C. (2013). Mystical, spiritual, and religious experiences. In R. F. Paloutzian, & C. L. Parks (Eds.), *Handbook of religion and spirituality* (2nd ed., pp. 422-440). New York, NY: Guilford Press.

Huntley, H. E. (1970). *The divine proportion*. New York, NY: Dover Publications.

Iosa, M., Fusco, A., Marchetti, F., Morone, G., Caltagirone, C., Paolucci, S., & Peppe, A. (2013). The golden ratio of gait harmony: Repetitive proportions of repetitive gait phases. *BioMed Research International*, *2013*. https://doi.org/10.1155/2013/918642

James, W. (1902/1985). *The varieties of religious experience*. Cambridge, MA: Harvard University Press.

Jung, C. G. (1952). Synchronicity: An acausal connecting principle. *The Collected Works of CG Jung, 8.*

Kail, R. V., & Cavanaugh, J. C. (2010). *Human development. A life span view*. Belmont, CA: Wadsworth.

Kelly, E. W., & Tucker, J. B. (2015). Research methods with spontaneous case studies. In J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 63-76). Jefferson, NC: McFarland.

Kelsey, T. W., Li, L. Q., Mitchell, R. T., Whelan, A., Anderson, R. A., & Wallace, W. H. B. (2014). A validated age-related normative model for male total testosterone shows increasing variance but no decline after age 40 years. *PLoS ONE*, *9*(10), e109346. https://doi.org/10.1371/journal.pone.0109346

Kvavilashvili, L., Kornbrot, D. E., Mash, V., Cockburn, J., & Milne, A. (2009). Differential effects of age on prospective and retrospective memory tasks in young, young-old, and old-old adults. *Memory*, *17*(2), 180-196. https://doi.org/10.1080/09658210802194366

Lagarkova, M. A., Eremeev, A. V., Svetlakov, A. V., Rubtsov, N. B., & Kiselev, S. L. (2010). Human embryonic stem cell lines isolation, cultivation, and characterization. *In Vitro Cellular & Developmental Biology-Animal*, *46*(3-4), 284-293. https://doi.org/10.1007/s11626-010-9282-6

Landers, R. N., Brusso, R. C., Cavanaugh, K. J., & Collmus, A. B. (2016). A primer on theory-driven web scraping: Automatic extraction of big data from the Internet for use in psychological research. *Psychological Methods*, *21*(4), 475-492. https://doi.org/10.1037/met0000081

Linage, G., Montoya, F., Sarmiento, A., Showalter, K., & Parmananda, P. (2006). Fibonacci order in the period-doubling cascade to chaos. *Physics Letters*, *A359*, 638-639. https://doi.org/10.1016/j.physleta.2006.07.036

Lindner, J. F., Kohar, V., Kia, B., Hippke, M., Learned, J. G., & Ditto, W. L. (2015). Strange nonchaotic stars. *Physical Review Letters*, *114*(5), 054101. https://doi.org/10.1103/PhysRevLett.114.054101

Löwdin, P. O. (1966). Quantum genetics and the aperiodic solid: Some aspects on the biological problems of heredity, mutations, aging, and tumors in view of the quantum theory of the DNA molecule. In P. O. Löwdin (Ed.), *Advances in Quantum Chemistry* (Vol. 2, pp. 213-360). New York, NY: Academic Press. https://doi.org/10.1016/s0065-3276(08)60076-3

Martin, F., Carminati, F., & Carminati, G. G. (2010). Quantum information, oscillations and the psyche. *Physics of Particles and Nuclei*, *41*(3), 425-451. https://doi.org/10.1134/S1063779610030032

Okabe, T. (2011). Physical phenomenology of phyllotaxis. *Journal of Theoretical Biology*, *280*(1), 63-75. https://doi.org/10.1016/j.jtbi.2011.03.037

Paul, S. P. (2016). Golden spirals and scalp whorls: nature's own design for rapid expansion. *PloS One*, *11*(9), e0162026. https://doi.org/10.1371/journal.pone.0162026

Pew Research Center. (2014). *Religious landscape survey*. Retrieved from http://www.pewforum.org/religious-landscape-study/

Piaget, J. (1970). Piagets theory. In P. H. Mussen (Ed.), *Carmichaels Handbook of Child Development* (pp. 703-732). New York, NY: Wiley.

Piedmont, R. L. (2005). The role of personality in understanding religious and spiritual constructs. In R. F. Paloutzian, & C. L. Park (Eds.), *Handbook of the psychology of religion and spirituality* (pp. 253-273). New York, NY: Guilford Press.

Ricketts, R. M. (1982). Divine proportion in facial esthetics. *Clinics in Plastic Surgery*, *9*(4), 401-422.

Roberts, G. J., Lucas, V. S., Andiappan, M., & McDonald, F. (2017). Dental age estimation: Pattern recognition of root canal widths of mandibular molars. A novel mandibular maturity marker at the 18-year threshold. *Journal of Forensic Sciences*, *62*(2), 351-354. https://doi.org/10.1111/1556-4029.13287

Rodgers, C. T., & Hore, P. J. (2009). Chemical magnetoreception in birds: The radical pair mechanism. *Proceedings of the National Academy of Sciences*, *106*(2), 353-360. https://doi.org/10.1073/pnas.0711968106

Rose, N. (1991). Design and development of wholeness: Waskom's paradigm. *The Educational Forum*, *55*, 243-259. https://doi.org/10.1080/00131729109335650

Russell, P. A. (2000). The aesthetics of rectangle proportion: Effects of judgment scale and context. *The American Journal of Psychology*, *113*(1), 27. https://doi.org/10.2307/1423459

Sacco, R. G. (2013). Re-envisaging the eight developmental stages of Erik Erikson: The Fibonacci Life-Chart Method (FLCM). *Journal of Educational and Developmental Psychology*, *3*(1), 140-146. https://doi.org/10.5539/jedp.v3n1p140

Sacco, R. G. (2016). The Fibonacci Life-Chart Method (FLCM) as a foundation for Carl Jung's theory of synchronicity. *Journal of Analytical Psychology*, *61*(2), 203-222. https://doi.org/10.1111/1468-5922.12204

Sarovar, M., Ishizaki, A., Fleming, G., & Whaley, K. (2010). Quantum entanglement in photosynthetic light-harvesting complexes. *Nature Physics*, *6*(6), 462-467. https://doi.org/10.1038/nphys1652

Smith, T. W. (2006). The national spiritual transformation study. *Journal for the Scientific Study of Religion, 45*(2), 283-296. https://doi.org/10.1111/j.1468-5906.2006.00306.x

Staff, L., Hurd, P., Reale, L., Seoighe, C., Rockwood, A., & Gehring, C. (2012). The hidden geometries of the Arabidopsis thaliana epidermis. *PloS One, 7*(9), e43546. https://doi.org/10.1371/journal.pone.0043546

Stapp, H. (2009). Nonlocality. In D. Greenberger, K. Hentschel, & F. Weinert (Eds.), *Compendium of Quantum Mechanics* (pp. 405-410). New York, NY: Springer. https://doi.org/10.1007/978-3-540-70626-7_125

Tehrani, F. R., Solaymani-Dodaran, M., Tohidi, M., Gohari, M. R., & Azizi, F. (2013). Modeling age at menopause using serum concentration of anti-mullerian hormone. *The Journal of Clinical Endocrinology & Metabolism, 98*(2), 729-735. https://doi.org/10.1210/jc.2012-3176

Terracciano, A., Costa, P. T., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin, 32*(8), 999-1009. https://doi.org/10.1177/0146167206288599

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of perception and action*. Cambridge, MA: MIT Press.

Turin, L. (2002). A method for the calculation of odor character from molecular structure. *Journal of Theoretical Biology, 216*(3), 367-385. https://doi.org/10.1006/jtbi.2001.2504

Underwood, L. G., & Teresi, J. A. (2002). The Daily Spiritual Experience Scale: Development, theoretical description, reliability, exploratory factor analysis, and preliminary construct validity using health-related data. *Annals of Behavioral Medicine, 24*, 22-33. https://doi.org/10.1207/S15324796ABM2401_04

Waskom, J. (1972). *The magic of design and other unpublished talks*. Loveland, Colorado.

Werdel, M. B., Dy-Liacco, G. S., Ciarrocchi, J. W., Wicks, R. J., & Breslford, G. M. (2014). The unique role of spirituality in the process of growth following stress and trauma. *Pastoral Psychology, 63*(1), 57-71. https://doi.org/10.1007/s11089-013-0538-4

Yamagishi, M. E. B., & Shimabukuro, A. I. (2008). Nucleotide frequencies in human genome and Fibonacci numbers. *Bulletin of Mathematical Biology, 70*(3), 643-653. https://doi.org/10.1007/s11538-007-9261-6

Zinnbauer, B. J., & Pargament, K. I. (2005). Religiousness and spirituality. In R. F. Paloutzian, & C. L. Park (Eds.), *Handbook of the psychology of religion and spirituality* (pp. 21-42). New York, NY: Guilford Press.

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Research on the Organizational Characteristics of Good High School Students' Mathematical Cognitive Structure Based on the Network Block-Modeling Analysis

Sun Dandan[1] & Yang Zezhong[1]

[1] The School of Mathematics&Statistics, Shandong Normal University, Jinan, China

Correspondence: Sun Dandan, The School of Mathematics&Statistics, Shandong Normal University, Jinan, China. Tel: 86-187-6617-3312. E-mail: 18766173312@163.com

Received: March 25, 2017 Accepted: March 31, 2017 Online Published: April 16, 2017

doi:10.5539/jedp.v7n2p14 URL: http://doi.org/10.5539/jedp.v7n2p14

Abstract

Through the direct detection and quantitative analysis of 44 concepts related to trigonometric functions in the mathematical cognitive structure of the 213 students in grade one in senior middle school, this paper finds that the mathematical cognitive structure of senior high school students has the following characteristics: In the cognitive structure of the mind, knowledge can be divided into different blocks according to the degree of relevance, the basis and the scale of the blocks, the degree of interaction in the block and between the blocks are not the same. The knowledge in a good cognitive structure should be organized in the form of blocks, and the blocks have a more scientific basis. The members in block have more obvious common features and each block is relatively large and covers more knowledge points. The block is closely linked, in addition, there must be a higher intensity of the link between the blocks, which can make the entire network of knowledge as a whole, and then it is conducive to the flow of information.

Keywords: mathematical cognitive structure, network analysis, organizational characteristics

1. Introduction

Mathematical cognitive structure is a internal structure with distinct individual characteristics, which is created by individuals and based on the studied mathematics knowledge combined with their own cognitive characteristics and non-intelligence factors and so on (Cao C. & Cai J., 1989; Li, 2001). Mathematical cognitive structure plays an important role in individual's mathematical activities. It not only affects the individual to understand and master the mathematical knowledge, but also affects the individual's application of mathematical knowledge, etc. (Yang, 1993; Yu, 2004). Therefore, in the actual teaching, the teachers expect to help the students to form a good cognitive structure of mathematics. To achieve this goal, there have been many researches on the characteristics of good mathematical cognitive structure in recent years, and achieved fruitful results.

Guan and He believed good mathematical cognitive structure should contain abundant basic knowledge of mathematics, especially a lot of knowledge about the production and the problem solving strategy (Guan, 1998; He, 2002). Wang, Han and Wang proposed that the characterization of ideas in good mathematical cognitive structure should be reasonable, clear, individualistic, interconnected, and flexible (Wang G. & Wang Y., 2004; Han & Wang, 2005). Yu thought that mathematical knowledge in good mathematical cognitive structure may organize together based on the knowledge field and knowledge system. The so-called knowledge field was a knowledge network, in which a large number of knowledge connected with each other according to the equivalence relations, the so-called knowledge system was a series of mathematical knowledge that was formed though abstract or deduction of some knowledge (Yu, 2011).

Review these studies, however, the vast majority are based on the thought of teaching activity experience combined with theory, there is little direct detection and quantitative research of students' mathematical cognitive structure. And, most of these studies concentrate on what content should the good mathematical cognitive structure has, little study is about how to organize the knowledge (Sun & Yang, 2015; Sun & Yang, 2015). The research about what content should be contained in good mathematical cognitive structure is

absolutely important, adopting the method of thinking is also very necessary, but, direct detection and quantitative research of good mathematical cognitive structure, especially of organizational form of knowledge in it is more important, that not only can make the study of good mathematical cognitive structure more in-depth, more clear, but also can make the training of good mathematical cognitive structure more targeted and directional in mathematics teaching, so the author choose part of senior grade one students and do a direct quantitative analysis of characteristics of their mathematical cognitive structure.

2. Methodology

2.1 Participants

We choose 213 students in grade one in senior middle school, who respectively come from two high schools in Jinan and two in Zibo of Shandong province as the research object. About the selection of top students, this study is based on the current international popular standard, namely the characteristics of usual mathematics learning and the score. The study of mathematical top students is mainly positive, efficient, and their method of study is reasonable, the performance is good and stable.

2.2 Methods

By applying the method of concept map and block-modeling, we conduct this study. The method of concept map is used to study the internal cognitive structure based on the concept diagram that individual drew, this method is commonly used to detect individual cognitive structure directly in psychology (Zhang & Chen, 2000). Block-modeling is one kind of the network analysis method, network analysis is just emerging in recent years, it is a kind of method that is used to analyze network based on the arithmetic of data relation (Liu, 2009). This method is widely used in studies about social network analysis early, and mainly used to study the density, centrality, relationship between subgroups etc. of the network to find out the characteristics of the network and the differences between different networks. Recently, this method is gradually being introduced to the network analysis related to the education and psychology (Luo, 2010). Block modeling is a kind of method used to study network location model, now also has been used in the study of organizational issues, and a large number of small group research, etc.

2.3 Instrument

We choose 44 mathematical concepts as study material, which come from the chapter named trigonometric function in high school mathematics textbook compulsory four published by people's education press (44 concepts are as shown in Table 1). We choose this part, because the characteristics of concepts in this part are obvious, the connections between concepts are numerous, and this part is also relatively concentrated in the present high school mathematics teaching.

Table 1. Concepts about trigonometric function

No.	concept	No.	concept	No.	concept	No.	concept
1	cyclic transformation	12	central angle	23	sinusoid	34	even function
2	acute angle	13	radian	24	phase	35	odd function
3	trigonometric function value	14	angle system	25	initial phase	36	minimal positive period
4	trigonometric function line	15	trigonometric function	26	cosine	37	period
5	tangent line	16	round angle	27	frequency	38	periodic function
6	cosine line	17	quadrant angle	28	amplitude	39	functional images
7	sine line	18	terminal side	29	domain of definition	40	induction formula
8	directed line segment	19	initial line	30	range	41	tangent function value
9	tangent function	20	zero angle	31	interval	42	cosine function value
10	cosine function	21	radian measure	32	decreasing function	43	obtuse angle
11	sine function	22	arbitrary angles.	33	increasing function	44	unit circle

2.4 Data Collection

First, let students draw the map of the relationships between these concepts on paper—that is, the concept map, and mark the compact degree of relationship that they think in the attachment between the adjacent two concepts. The compact degree of relationship is respectively represented with an integer between 1 to 5 from low to high. Then transform the concept map that students draw into one-mode multi-valued relationship matrix in prepare for the analysis of network.

2.5 Data Analysis

First, use the network analysis software—Ucinet6.0 and NetDraw to deal with the multi-valued relationship matrix, then compare the situations of top students, medium and general students.

3. Process

This study is carried out when the high school students have learned the chapter named trigonometric function in compulsory four. In order to collect the data effectively and easily, the researchers take the way of cluster sampling. In the entire process, the experimenters is always acted by the researchers and mathematics teacher. The process of data collection is as follows: first of all, let the students recall the content in the chapter of trigonometric function, and then present the 44 math concepts that just studied to the students, in order to help students have a more clear understanding of these concepts. Later, let the students draw the map of relationships between these concepts according to their own understanding.

There are 213 students in the four classes in grade one in two senior middle schools, including 96 boys and 117 girls, 31 top students, 127 medium students, and 55 general students. This study finally receive 213 concept maps, 7 sets are rejected because they don't meet the requirements (mainly because the picture is not clear and disorderly), we get 206 concept maps that are suitable for further analysis. For the 206 concept maps, the researchers transform them into a one-mode multi-valued relationship matrix first, and carry on the preliminary sorting depending on the type of the level of the students, then analyze the block modeling of concept networks using network software Ucinet6.0. The materials of middle and general students are also surveyed and analyzed here, the reason is that we want put their data as compare objects, so that we can analyze the characteristics of top students' mathematical cognitive structure more in-depth.

4. Results

In order to analysis the structure of knowledge network, for different types of knowledge networks, at first, we calculate the number of subgroups (locations) in the network and then show density and image matrix between different positions, at last show the simplified diagram of the relationship. It can be clearly seen that the knowledge network is divided into a number of "gangs" and how they are connected with each other. For this purpose, the three kinds of mean value matrix of the top students, the medium students and the general students are treated symmetrically, and the tree diagrams of the network are divided into 8 blocks by the Concor method.

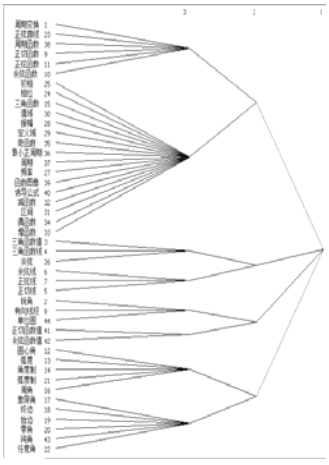


Figure 1. Top students

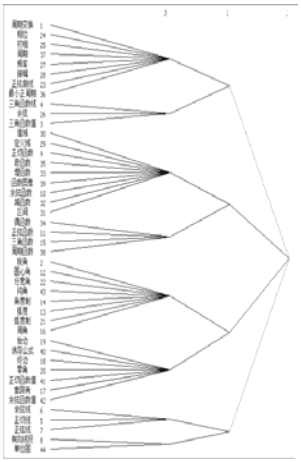


Figure 2. Medium students

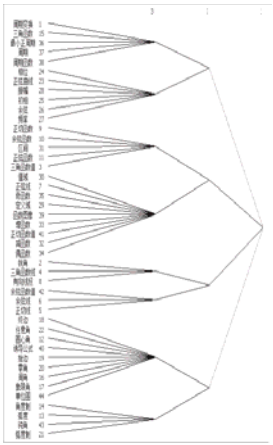


Figure 3. General students

It can be seen from the above three diagrams, whether it is condensed into four subgroups or eight subgroups, the above 44 mathematical concepts can be divided into a large subgroup and a few subgroups in the cognitive structure of the top students, instead, in the cognitive structure of medium students and the general students, they are basically the small subgroups with relatively average number of elements, without large subgroup with obvious numerous elements. This situation is more obvious with the increasing of cohesive subgroups.

When condensed into eight subgroups, the condensation condition of the above 44 mathematical concepts in the top students', medium students' and the general students' cognitive structure is very different. The common features of the majority of subgroups in the cognitive structure of top students are more obvious, and the relationship between each other is more clear and reasonable. For example, the concepts of the top students' second subgroup are basically about the basic properties of trigonometric functions. The concepts of the fourth subgroup of the top students is closely related to the trigonometric function line. The concepts of the sixth subgroup are obviously related to the trigonometric function value. The concepts of the seventh subgroup are related to the radian system. The concepts of the eighth subgroup are related to arbitrary angle. Moreover, when the whole network is divided into 4 blocks, the first and the second subgroups of the top students will condense into a group further which contain all the key concepts of the basic properties of trigonometric functions. In detail, the third and the fourth subgroups can be regarded as the expansion of trigonometric function line; the seventh and the eighth subgroups are all based on the comparative basis of arbitrary angles and radians. Using the same analytical thinking to look backward the subgroup of medium students and general students, we can find some general and common features. For example, the concepts of the first subgroup of medium students are related to the properties of trigonometric functions; the concepts of the seventh and the eighth subgroups are

related to the trigonometric function line; the concepts of the second subgroup of general students are related to the properties of trigonometric functions. However, from the perspective of each subgroup's members of medium students and general students, it's not very easy to find common points because many subgroups are associated with other uncorrelated concepts. For example, the seventh subgroup of the general students is almost about arbitrary angle, but also doping induction formula and unit circle. Secondly, from a macro point of view, there is a very obvious and important difference among the top, medium and general students. The property of the trigonometric function in this chapter is a very important knowledge point. In the cognitive structure of top students, the unique properties of trigonometric functions such as initial phase, phase and amplitude can be combined with the properties of the functions previously studied, such as monotonicity and parity and then form an organic whole and reflected in the second subgroup. However, these two parts of properties in the cognitive structure of medium students and general students scatter in three subgroups. Such knowledge organization is obviously not so scientific as that of the top students which link the previous and following knowledge closely according to a certain standard of knowledge. It is worth noting that in the third subgroup of medium students, although it is mainly about general function there are two kinds of trigonometric functions, which shows that there is intention to contact the two parts' knowledge. However, this kind of connection is weaker in the subgroups of general students. At this time, the density matrix of top students, medium students and general students is shown in Tables 2, 3, 4.

Table 2. The density matrix of top students

No.	1	2	3	4	5	6	7	8
1	1.11	1.12	0.46	0.46	0.07	0.20	0.04	0.01
2	1.12	0.31	0.10	0.24	0.05	0.22	0.03	0.02
3	0.46	0.10	0.85	1.70	0.49	0.15	0.03	0.08
4	0.46	0.24	1.70	0.72	1.23	0.27	0.00	0.03
5	0.07	0.05	0.49	1.23	0.23	0.00	0.26	0.41
6	0.20	0.22	0.15	0.27	0.00	0.75	0.11	0.38
7	0.04	0.03	0.03	0.00	0.26	0.11	1.21	0.33
8	0.01	0.02	0.08	0.03	0.41	0.38	0.33	1.11

Table 3. The density matrix of medium students

No.	1	2	3	4	5	6	7	8
1	0.36	0.21	0.12	0.26	0.02	0.01	0.06	0.01
2	0.21	0.73	0.15	0.19	0.05	0.09	0.54	0.08
3	0.12	0.15	0.67	0.48	0.01	0.05	0.04	0.00
4	0.26	0.19	0.48	0.46	0.07	0.10	0.07	0.08
5	0.02	0.05	0.01	0.07	0.42	0.14	0.02	0.15
6	0.01	0.09	0.05	0.10	0.14	0.34	0.03	0.09
7	0.06	0.54	0.04	0.07	0.02	0.03	0.64	0.67
8	0.01	0.08	0.00	0.08	0.15	0.09	0.67	0.00

Table 4. The density matrix of general students

No.	1	2	3	4	5	6	7	8
1	0.65	0.23	0.21	0.08	0.01	0.03	0.02	0.00
2	0.23	0.24	0.09	0.02	0.00	0.03	0.00	0.00
3	0.21	0.09	0.06	0.36	0.07	0.20	0.04	0.00
4	0.08	0.02	0.36	0.10	0.06	0.06	0.02	0.00
5	0.01	0.00	0.07	0.06	0.17	0.44	0.03	0.05
6	0.03	0.03	0.20	0.06	0.44	0.14	0.08	0.00
7	0.02	0.00	0.04	0.02	0.03	0.08	0.30	0.12
8	0.00	0.00	0.00	0.00	0.05	0.00	0.12	0.43

The density matrix is a symmetric matrix and the number in the diagonal position represents the density of each subgroup, the number in the remaining position represents the correlation density of corresponding rows and columns of subgroups. It can be seen from Tables 2, 3, 4, the density of the cohesive subgroups of the top students is range from 0 to 1.7, with an average about 0.36. The density of the cohesive subgroups of the medium students is range from 0 to 0.728, with an average about 0.17. The density of the cohesive subgroups of the general students is range form 0 to 0.65, with an average about 0.10. No matter in terms of the maximum density or the average density, the top students are both higher than the medium students, the medium students are higher than the general students. This shows that, on the whole, the degree of connection relationship between different subgroups in the cognitive structure of the top students is generally greater than that of the medium students. On the basis of the relationship in and between the subgroups of the cognitive structure network, the top students are better than the medium students and the medium students are better than the general students.

It can be also seen from Tables 2, 3, 4, that the maximum degree of closeness between the different subgroups in the cognitive structure of the top students is about 1.7—the third subgroup and the fourth subgroup. However, the maximum degree of connection between different subgroups in the cognitive structure of medium students is only 0.667—the seventh subgroup and the eighth subgroup. The maximum degree of connection between different subgroups in the cognitive structure of general students is 0.444—the fifth subgroup and the sixth subgroup. Interestingly, the maximum degree of closeness of top students, medium students, general students both appears between two subgroups trigonometric—function line, directed line segment and sine line, cosine line, tangent line, which is also a scientific fact that is in accord with objective conditions.

The densities of the multi valued symmetric matrix networks of the general students, the medium students and the top students are respectively 0.0928, 0.1533, 0.3279. Regard the density of three types of network as the critical value. Then, the density of the corresponding density matrix higher than the critical value is recorded as 1, and the density less than the critical value is recorded as 0. Finally, we get the image matrix of general students, medium students and top students and the simplified relation diagram can be got according to it.

Image matrix is a two numerical value symmetric matrix. When the numerical value is 1, it indicates that there is a connection between corresponding block in the sense of higher than average density. When the numerical value is 0, there is no connections. Each circle in the simplified diagram represents a subgroup, curves with arrows on each circle represents the relationship from the circle “issued” back to the circle, that is, the density of the subgroup is higher than the overall matrix density. If there is a connection between circles, which means in the sense of higher than average density there is a “communication” between subgroups. If there is no connection between points that means “no communication”.

Table 5. Image matrix of top students

No.	1	2	3	4	5	6	7	8
1	1	1	1	1	0	0	0	0
2	1	0	0	0	0	0	0	0
3	1	0	1	1	1	0	0	0
4	1	0	1	1	1	0	0	0
5	0	0	1	1	0	0	0	1
6	0	0	0	0	0	1	0	1
7	0	0	0	0	0	0	1	1
8	0	0	0	0	1	1	1	1

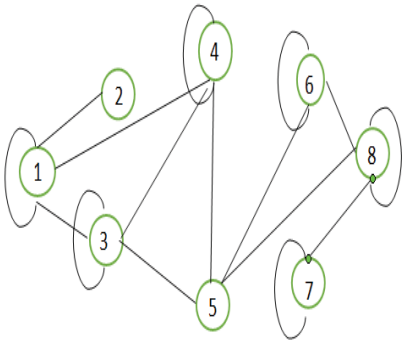


Figure 4. Simplified diagram of top students

Table 6. Image matrix of medium students

No.	1	2	3	4	5	6	7	8
1	1	1	0	1	0	0	0	0
2	1	1	0	1	0	0	1	0
3	0	0	1	1	0	0	0	0
4	1	1	1	1	0	0	0	0
5	0	0	0	0	1	0	0	0
6	0	0	0	0	0	1	0	0
7	0	1	0	0	0	0	1	1
8	0	0	0	0	0	0	1	0

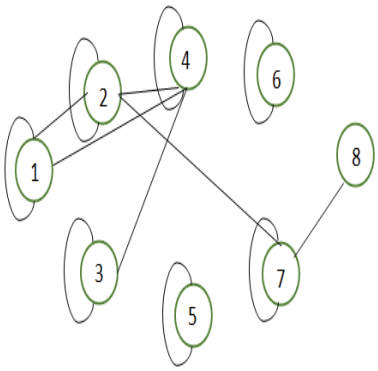


Figure 5. Simplified diagram of medium students

Table 7. Image matrix of medium students

N	1	2	3	4	5	6	7	8
1	1	1	1	0	0	0	0	0
2	1	1	0	0	0	0	0	0
3	1	0	0	1	0	1	0	0
4	0	0	1	1	0	0	0	0
5	0	0	0	0	1	1	0	0
6	0	0	1	0	1	1	0	0
7	0	0	0	0	0	0	1	1
8	0	0	0	0	0	0	1	1

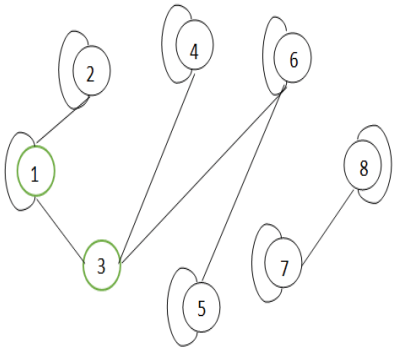


Figure 6. Simplified diagram of general students

It can be seen clearly that there are no scattered subgroups in the concept network of top students. In details, there is a close connection between the 8 positions. The fifth and the sixth subgroups of medium students are isolated and the seventh and the eighth subgroups of general students are not associated with other subgroups. The relatively isolated subgroups of the concept network of medium and general students are arbitrary angles and radians. In addition, there is mutual relation in the simplified relationship diagram of top students and medium students. In this case, three small groups may form a relatively large group which is conducive to the circulation of information, whereas a similar situation does not appear in the simplified concept diagrams of general students. The relationship between the subgroups of general students is basically linear, and there is no greater subgroup.

4. Conclusions and Suggestion

In the cognitive structure of the mind, knowledge can be divided into different blocks according to the degree of relevance. The basis and the scale of the blocks, the degree of interaction in the block and between the blocks are not the same. In the knowledge structure of the top students, there are often a large block and a few small blocks when the knowledge is divided, nevertheless the members of the blocks of the medium students and the general students have relatively little difference. In details, the top students may embeds the specific knowledge in the higher level of knowledge, thus, a more general relation between knowledge is established, and a large knowledge block is obtained. The classification of the blocks of top students is more scientific, and the commonness of the members in the blocks is more clearer. In addition, the density in the block and between the blocks is relatively large, that is, the connection of the knowledge in the block and between the blocks is the most compact. The knowledge in a good cognitive structure should be organized in the form of blocks, and the blocks have a more scientific basis. The members in block have more obvious common features and each block is relatively large and covers more knowledge points. The block is closely linked, in addition, there must be a higher intensity of the link between the blocks, which can make the entire network of knowledge as a whole, and then it is conducive to the flow of information.

Based on this, in the teaching process, first of all, teachers should guide students to build knowledge blocks, that is, to teach clearly about the relationship between knowledge so that the knowledge scattered in the cognitive structure of students can be made into a block which has great benefit to students' knowledge memory, extraction and operation. In addition, teachers should not only pay attention to guide students to establish a clear knowledge block, but also strive to guide students to expand the scale of knowledge blocks. How to expand it? Starting from the edge knowledge points of the established knowledge blocks, we can establish the further connection between these knowledge points and other concepts, so as to expand the scale of knowledge blocks.

Acknowledgments

Supported by the project of research on organizational characteristic of mathematical cognitive based on the network analysis (SCX201614).

References

Cao, C. H., & Cai, J. F. (1989). *Introduction to Mathematics Education* (pp. 52-53). Nanjing: Jiangsu Education Press.

Guan, P. (1998). Principles of Cognitive Psychology on Forming Good Mathematical Cognitive Structure. *Theory and Practice of Education*, *18*(2), 40-45.

Han, B., & Wang, G. M. (2005). Thinking about the Role of Cognitive Structure in the Process of Problem Solving. *Journal of Middle School Mathematics (High School)*, *6*, 5-7.

He, X. Y. (2002). Instructional Strategies on Constructing well Mathematical Cognitive Structure. *Journal of Mathematics Education*, *11*(1), 24-27.

Li, S. Q. (2001). *PME: Psychology of Mathematics Education* (pp. 24-25). Shanghai: East China Normal University Press.

Liu, J. (2009). *Lectures on Whole Network Approach* (pp. 17-38). Shanghai: Gezhi Press.

Luo, J. D. (2010). *Social Network Analysis* (pp. 47-57). Beijing: Social Sciences Academic Press.

Sun, D. D., & Yang, Z. Z. (2015). Research on Good Mathematical Cognitive Structure in Mainland China. *Advances in Social and Behavioral Sciences: Proceeding of 2015 3rd Asian Conference on the Social Sciences*, *15*, 216-219.

Sun, D. D., & Yang, Z. Z. (2015). Study on Content and Organization of Mathematical Cognitive Structure in Mainland China. *Advances in Intelligent Systems Research*, *125*, 1521-1525. https://doi.org/10.2991/meici-15.2015.283

Wang, G. M., & Wang, Y. (2004). The comparison of the mathematical cognitive structure between the top and ordinary student in senior high school, the possible reason and teaching suggestions. *Preference for Middle School Mathematical Education*, *12*, 1-4.

Yang, Q. (1993). An Study on The Influence of Cognitive Structure to Mathematics Learning—One of the Explores about the Psychological Factors That Influences the Mathematics Learning. *Journal of Mathematics Education*, *2*(1), 66-70.

Yu, P. (2004). Influence of Self-controlling Ability and CPFS Structure on the Mathematical Achievement in High School Students. *Journal of Mathematics Education*, *13*(1), 23-26.

Yu, P. (2011). *China's Psychology Research of Mathematics Education in 30 Years*. Beijing: Science Press.

Zhang, J. W., & Chen, Q. (2000). Test Method of Cognitive Structure. *Journal of Psychological Science*, *6*, 750-751.

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Correlates of a University Counseling Center's Perceived Service Promptness

Gary Blau[1], Peter A DeMaria, Jr[2] & John DiMino[2]

[1] Human Resource Management Department, Fox School of Business, Temple University, Philadelphia, USA

[2] Tuttleman Counseling Services, Temple University, Philadelphia, USA

Correspondence: Gary Blau, Human Resource Management Department, Temple University, 349 Alter Hall, 1810 Liacouras Walk, Philadelphia, PA, 19122, USA. E-mail: gblau@temple.edu

Received: April 15, 2017 Accepted: May 8, 2017 Online Published: May 22, 2017

doi:10.5539/jedp.v7n2p24 URL: http://doi.org/10.5539/jedp.v7n2p24

Abstract

The overall goal of this study was to explore the usefulness of a Perceived Service Promptness (PSP) measure for University Counseling Centers (UCCs). As UCCs compete for university resources, helping a client as quickly as possible or PSP can help to support a UCC's documented ability to meet increasing client demands. Since no prior empirical research was found measuring PSP at a UCC, a four-item measure, adapted from a more general quality of service scale, was used. From August 2014 to May 2016, one hundred and seventeen non-urgent undergraduate students seeking counseling services filled out an online survey measuring demographics, client perceptions, wait measures, PSP and recommending the university. Confirmatory factor analysis and scale reliability data psychometrically supported the PSP scale. Correlational analyses showed that both wait time and wait bother experience were each significantly negatively related to PSP. However, hierarchical regression analyses showed that wait bother experience, but not wait time, significantly explained PSP beyond prior controlled-for demographic and client perception variables. In addition, PSP positively explained recommending the university beyond demographic, client perception and wait measures. Research limitations and future research issues are discussed.

Keywords: university counseling center, perceived service promptness, wait experience, recommend university

1. Introduction

There has been a dramatic increase in college student demand for University Counseling Center (UCC) services (Gallagher, 2014). Brunner, Wallace, Reymann, Sellers and McCabe (2014, p. 264) noted that this may be the "most stressed" generation of college students. A study of stress tolerance in a sample of millennial college students (Bland, Melton, Welle, & Bigham, 2012) showed that of the 29 lifestyle habits and coping skills assessed, only one, i.e., felt supported by family, friends and teachers, was a significant factor for helping high stress tolerance. A qualitative study showed that university professionals perceived that millennial college students depended on others for problem solving (Much, Wagener, Breitkreuetz, & Hellenbrand, 2014). A UCC can represent an important base of at least temporary support for millennial college students (Brunner et al., 2014).

Many UCCs have adopted a triage system to help meet this increased demand (Rockland-Miller & Eells, 2006; Shaffer et al., 2017). In a triage system, clients receive an initial screening evaluation and those classified as urgent (e.g., a danger to self or others, severe or psychotic symptoms) are given priority appointments and seen quickly. Non-urgent clients may have to wait to begin counseling. Research has found that a longer wait time for non-urgent clients led to higher subsequent no-show rate for beginning counseling (DiMino & Blau, 2012). What variables affect the relationship between wait time and showing up to begin counseling for non-urgent clients is in need of further study. One variable may be Perceived Service Promptness (PSP) of the UCC. In other words, does a non-urgent client perceive that the UCC wants to help him/her as quickly as possible? The overall goal of this study was to explore the usefulness of a PSP measure for UCCs.

While trying to meet increased service demand, UCCs need to compete with other campus service stakeholders for resources (Brunner et al., 2014; Hunt, Watkins, & Eisenberg, 2012). In the current context of increased

accountability and restricted budgets, research can help to document a UCC's ability to meet increased service demand. Valotta (2012) noted that client assessment measures are needed to help with planning UCC staffing and budgeting. Measuring PSP of a UCC can assist with such planning efforts. The positive responsiveness of a UCC, as 24/7 available service for students, can also have positive implications for a university's public relations and student recruitment (Bishop, 2010). The purposes of this study are to: (1) test the adaptability of a general service responsiveness scale for measuring UCC PSP; (2) investigate client perception and wait measure correlates of PSP; and (3) test if PSP can significantly explain recommending a university beyond controlled-for demographic, client perception and wait measure variables. Given the lack of prior research, these three issues are framed as separate research questions.

1.1 Measuring UCC PSP

PSP falls within the larger construct of service quality (Parasuraman, Zeithaml, & Berry, 1988). Quality of service across organizations has been successfully measured at the general consumer level using SERVQUAL, a 22-item measure broken down into five dimensions: tangibles, reliability, responsiveness, assurance, and empathy (Parasuraman et al., 1988). Responsiveness in SERVQUAL is measured by four items focusing on perceived promptness (helping as quickly as possible) in receiving service. Given the absence of an existing PSP measure applied to UCCs, these four items were adapted from SERVQUAL to measure PSP in this study.

Focusing on student samples in general health care settings, there has been limited research on PSP. One study measuring student perceptions of service quality at a student health center (Canel & Anderson Fletcher, 2001) found there was a large gap between higher student expectations versus the perceptions of health center employees' "willingness to help". A second student healthcare setting study found that perceived promptness was positively related to student satisfaction (Senic & Marinkovic, 2013). A recent qualitative study (Randall & Bewick, 2016) looked at counselor perceptions in the United Kingdom for improving service quality to students. Work style changes, including improved recording/note keeping processes, resulted in a more efficient service that was perceived to be more acceptable to students (Randall & Bewick, 2016). However, no prior empirical research was found specifically measuring UCC PSP. This leads to the first research question:

RQ1-Can an existing general service responsiveness scale can be adapted to adequately measure UCC PSP?

1.2 Antecedents of UCC PSP

Level of mental distress can motivate a student to seek counseling (Brunner, Wallace, Reymann, Sellers, & McCabe, 2014). However, perceived stigma for receiving psychological help can prevent a student from seeking counseling (Golberstein, Eisenberg, & Gollust, 2008). Making students more aware of and comfortable with seeking campus mental health services, such as reducing the stigma of receiving psychological help, should be part of a university's academic mission (Mowbray et al., 2006). For hesitant students, a longer wait time may serve as a roadblock to seeking services.

In a study with a very different patient population, longer perceived waiting time led to lower patient satisfaction for radiological outpatients across 6 hospitals in Belgium (Bielen & Demoulin, 2007). In a study with non-urgent undergraduate clients at a UCC, clients who reported waiting less than 2 weeks for their intake appointment perceived a lower stigma for receiving psychological help, stronger institutional commitment and a higher willingness to recommend the university to others, versus clients waiting more than 2 weeks (Blau et al., 2015). UCCs typically measure client wait time as part of their service assessment process (Shaffer et al., 2017). In addition to amount of waiting time, "wait experience" may be important to assess. Dissatisfaction with the wait experience (wait bother) led to lower perceived service quality for customers of a university cafeteria (Lee & Lambert, 2005). The above-cited research collectively suggests that perceived stigma for counseling, level of mental distress, wait time and wait bother experience may negatively affect PSP. This suggests the second research question:

RQ2-Will perceived stigma for counseling, level of mental distress, wait time and wait bother experience each be negatively related to PSP?

1.3 UCC PSP Related to Recommending a University

A UCC is generally considered to be part of campus support services, which has been found to predict student satisfaction (Elliott & Healy, 2001). Bielen and Demoulin (2007) found that higher patient satisfaction led to a higher likelihood of recommending the radiological services to relatives. Bishop (2010) noted that UCCs may positively contribute to the general recruitment of prospective students. UCC clients who have higher PSP may be more likely to recommend their university to others. For this relationship, PSP is being used as an

independent variable. Prior research (e.g., Winterrowd, Priniski, Achter, & Abhold, 2016) investigating outcomes of UCC counseling has recommended controlling for demographic variables (e.g., gender, ethnic background, age, status entering university). Controlling for demographic and other study variables first will allow for a stronger assessment of PSP impact. This suggests the third research question:

RQ3-Will PSP positively impact recommending a university beyond controlled-for demographic, client perception and wait measure variables?

2. Methods

2.1 Participants and Procedure

The study was conducted at a large urban public university in the Mid-Atlantic region of the United States. From August 2014 to May 2016, 117 undergraduate students who presented at the UCC (Tuttleman Counseling Services) seeking counseling services and were determined to be non-urgent (i.e., did not report suicidality or risk to self/others), completed the survey questionnaire after triage and before intake. Institutional Review Board (IRB) approval was given for all data collection. Survey responses went directly into a Qualtrics' data base. Consistent with recent national surveys, students indicated that anxiety (42%) and depression (27%) were the main reasons for coming to the UCC (Brunner et al., 2014).

2.2 Measures

The survey consisted of five measurement sections: (1) respondent demographics; (2) client perceptions; (3) wait measures; (4) perceived service promptness, and (5) recommending the university. Unless otherwise noted below, items used a seven-point Likert response scale, from 1 = strongly disagree to 7 = strongly agree. A small pilot test of item content for survey measures was done among four UCC clinicians and 10 undergraduate student interns at the UCC's self-help center. The clinicians' feedback qualitatively supported the content validity of the items (Nunnally, 1978), while the interns found the item language understandable for respondents. Cronbach's alpha for estimating the internal consistency reliability for each multi-item measure is reported.

2.2.1 Respondent Demographics

Gender, ethnic background, status entering university, current residential status, referral source for counseling (i.e., self-versus-other), and age were measured. A demographic breakdown of the UCC client sample showed: 69% female students; 73% white versus 27% non-white; 74% percent were non-transfer students versus 26% transfer students; 21% percent were commuter versus 79% were residential students; and 59% were self-referred with the remaining 41% were referred by others (e.g., family, university staff) to the UCC. Mean age was 20 years old. By comparison, for the fall of 2016, based on 28,203 matriculated undergraduates at the university: 53% were female, 57% were white, and 65% were non-transfer, and the mean age was 22 years old. Thus in terms of representativeness, the UCC client sample has higher percentages of female, white, and non-transfer students who are generally younger. However, the UCC client sample demographics are more consistent with 2014 data from the Association for University and College Counseling Directors Annual Survey (Reetz, Krylowicz, & Mistler, 2014) on the percentages of female (65%) and white students (67%) served by UCCs.

2.2.2 Client Perceptions

There are two measures: stigma for counseling and level of mental distress. Stigma for counseling was measured using a five-item scale (Komiya, Good, & Sherrod, 2000). A sample item is: "it is a sign of personal weakness or inadequacy to see a psychologist for emotional or interpersonal issues". Cronbach's alpha was .85. Level of mental distress was measured using one-item: "Please rate your current level of distress for the mental issue that brought you to counseling". An eight-point response scale used where: 1 = none, 2 = very little, 3 = a little, 4 = moderate, 5 = a lot, 6 = severe, 7 = very severe and 8 = unbearable.

2.2.3 Wait Measures

There are two measures, wait time and wait bother experience. Wait time was measured by asking "indicate as best you can how many days you waited between your initial screening or 'walk in' session (triage) and starting to see your assigned counselor (your intake session)". Wait bother experience was measured using one item asking "how you felt about waiting to start with your assigned counselor". A four-point response scale was used where: 1 = the waiting did not bother me at all, 2 = the waiting bothered me a little bit; 3 = the waiting bothered me moderately and 4 = the waiting bothered me a lot.

2.2.4 Perceived Service Promptness (PSP).

These four items were adapted from the four-item responsiveness measure of SERVQUAL (Parasuraman et al., 1988). Parasuraman et al. (1988) found a Cronbach's alpha of .82, with four-items, using a sample of 200 adult service-user respondents. Parasurman et al. (1988, p. 23) gave the following definition for responsiveness: "willingness to help customers and provide prompt service". This definition was operationalized using the following four items: "I received prompt service from the UCC"; "The UCC told me when my counseling would begin"; "The UCC wanted to help me as quickly as possible"; and "The UCC was not too busy to respond promptly to me". As noted earlier, the response scale for each item was 1 = strongly disagree to 7 = strongly agree. This data is analyzed to test RQ1.

2.2.5 Recommend University

This three-item measure was previously used (Blau et al., 2015). A sample item is: "I would recommend this university to my family and friends". Cronbach's alpha was .92.

2.3 Data Analysis

Confirmatory Factor Analysis (CFA) was used to test the first Research Question (RQ1). For RQ2 and RQ3, with a sample size n = 117, significance level of .05, assuming a medium effect size correlation of .30, the statistical power was .91 (Cohen, 1969). For RQ2 and RQ3, stepwise regression analyses (Stevens, 1996) were used to evaluate the impact of demographic variables (Step 1), then the incremental impact of client perception variables (Step 2), and then wait measures (Step 3) in three separate models on PSP. In these first three (steps) models, PSP was used as a dependent variable. Binary measures for the demographic variables allowed them to be directly entered in the regression models (Stevens, 1996). Assuming an anticipated effect size of .15 with 11 predictors using a probability level of .05, with the current sample size, the statistical power was .80. In the final regression model (Model 4), PSP is used as an independent variable, and the impact of PSP for significantly explaining recommend the university after controlling for demographic, client perception and wait measure variables was tested. For the regression analyses it was determined that the assumptions of no multicollinearity, linearity, and homoscedasticity were satisfactorily met (Stevens, 1996). Care was taken to avoid four common errors associated with hierarchical/stepwise regression: neglect of a theoretical basis for using; violation of causal priority; use in an exploratory manner; and misinterpretation of results (Petrocelli, 2003).

3. Results

3.1 Test of RQ1

Using CFA, the following fit statistics of the four-items to the PSP construct were found: X^2 (2, N = 117) = 2.38, p > .05; Adjusted Goodness of Fit (AGFI) = .96; Comparative Fit Index (CFI) = .99; Root Mean Square Residual (RMR) = .03; and Root Mean Square Error of Approximation (RMSEA) = .04. Thresholds for acceptable fit (Bentler, 1990) should be at least .90 (AGFI, CFI) and less than .08 for error measures (RMR, RMSEA). Prior work has used similar CFA statistics to support recent scales, i.e., Social Identities and Attitudes Scale (Smith & Cokley, 2016), and College and Career Readiness Counseling Support (Lapan, Poynton, Marcotte, Marland, & Milam, 2017). Cronbach's alpha for the four-item PSP scale was .79. Each item contributed to the scale reliability, and the average inter-item correlation was r = .46. Collectively these results found a psychometrically sound four-item scale for measuring PSP, supporting RQ1.

3.2 Tests of RQ2 and RQ3

To make interpretation easier, the aggregate score for a multi-item measure was divided by the number of items so that the mean (M) was based on the measure's response scale. The means, standard deviations and correlations between continuous variables are reported in Table 1. The mean level for PSP was 5.89 out of 7, which indicates a generally high level. Level of mental distress ($\underline{r}$ = -.20), wait time ($\underline{r}$ = -.35) and wait bother experience ($\underline{r}$ = -.48) were each significantly negatively related to PSP. However, perceived stigma for counseling was not related to PSP ($\underline{r}$ = -.05). This indicates partial support for RQ2. Based on translating r^2 partial eta squared into Cohen's effect size (Cohen, 1969), level of mental distress is a medium effect size while wait time and wait bother experience are large effect sizes. Finally, PSP was positively related to recommending the university ($\underline{r}$ = .23), which is a medium effect size.

Table 1. Means, standard deviations, reliabilities and correlations for continuous variables

Measure	M	SD	1	2	3	4	5	6	7
1. Age[a]	4.09	2.37	(NA)[f]						
2. Stigma for Counseling[b]	3.39	1.39	-.13	(.85)					
3. Level of Mental Distress[c]	5.00	1.08	.08	.13	(NA)				
4. Wait Time[d]	14.84	7.67	.12	-.12	.01	(NA)			
5. Wait Bother Experience[e]	2.05	.91	.17	.18	.20*	.40**	(NA)		
6. Perceived Service Promptness[b]	5.89	1.03	.01	-.05	-.20*	-.35**	-.48**	(.79)	
7. Recommend University[b]	5.79	1.25	-.09	-.11	-.25**	-.01	-.15	.23*	(.92)

N = 117, * p < .05; ** p < .01 (two-tailed).

[a] Age, 1 = less than 18 years old to 14 = at least 30 years old, so 2 = 18 years old, 3 = 19 years old, 4 = 20 years old, 5 = 21 years old, 6 = 22 years old……14 = at least 30 years old.

[b] Stigma for Counseling, Perceived Service Promptness, and Recommend University, 1 = strongly disagree to 7 = strongly agree.

[c] Level of Mental Distress, 1 = none to 8 = unbearable.

[d] Wait Time, 1 = 1 day to 41 = over 40 days.

[e] Wait Bother Experience, 1 = waiting did not bother me at all to 4 = waiting bothered me a lot.

[f] (Coefficient alpha), NA = not applicable.

Table 2 shows the stepwise regression model results for testing the RQ2 and RQ3. For RQ2, Models 1, 2 and 3 evaluate the impact of variables for explaining PSP as the dependent variable. Looking at the impact of variables for explaining PSP first, the initial significance of several demographic variables in Models 1 and 2, i.e., being a transfer student, commuter and self-referral, and level of mental distress (Model 2 only) impacting PSP decrease to non-significance in Model 3. In Model 3 only wait bother experience (b = -.427) significantly negatively affects PSP, but wait time does not (b = -.110). The results of Model 3 provide very limited support for RQ2. Model 4 evaluates the impact of adding in PSP, as an independent variable, to explain recommending the university as the dependent variable, beyond the demographic, client perceptions and wait measure variables. Model 4 finds that being a non-commuter (b = -.252), and PSP (b = .264), each significantly increase recommending the university. PSP significantly explains additional significant variance (R^2 = 4.4%) in recommending the university beyond controlled for variables. This result supports RQ3. Overall, 35% of the variance (R^2) was explained for PSP and 17% for recommending the university.

Table 2. Results of stepwise regression analyses on perceived service promptness and recommend university

Dependent Variable	Perceived Service Promptness			Recommend University
Independent Variable	Model 1[f]	Model 2[f]	Model 3[f]	Model 4[g]
Step 1-Demographics				
Gender[a]	-.008	-.014	.112	.047
Ethnic Background[b]	.031	-.018	-.042	-.027
Status Entering University[c]	.234*	.228*	.136	-.024
Current Residential Status[d]	.184	.195*	.137	-.252*
Referral Source for Counseling[e]	-.179	-.194*	-.138	.056
Age	-.145	-.126	.040	-.020
Step 2-Client Perceptions				
Stigma for Counseling		.024	.090	-.095
Level of Mental Distress		-.226*	-.154	-.146

Step 3-Wait Measures				
Wait Time			-.110	.052
Wait Bother Experience			-.427**	-.039
Step 4				
Perceived Service Promptness				.264*
F	2.57*	2.76*	5.73**	1.94*
Degrees of freedom	6,110	8,108	10,106	11,105
R^2	.123	.170	.351	.172
Adjusted R^2	.075	.108	.290	.093
Change in R^2		.047*	.181**	.044*

Note. Standardized regression coefficients reported. N = 117.

* p < .05; ** p < .01 (two-tailed).

[a] gender (1 = male, 2 = female); [b] ethnic background (1 = non-white, 2 = white); [c] status entering university (1 = non-transfer, 2 = transfer); [d] current residential status (1 = non-commuter, 2 = commuter); [e] referral source for counseling (1 = self, 2 = other).

[f] for Models1, 2 and 3, Perceived Service Promptness is used as the dependent variable.

[g] for Model 4, all correlates for explaining Intent to Recommend are entered stepwise and the incremental variance of Perceived Service Promptness in Step 4, as an independent variable, is evaluated for incrementally explaining Recommend University.

4. Discussion

Kitzrow (2009, p. 657) noted the importance of college student mental health issues "as an institutional responsibility and priority". With the demand for UCC services increasing (Gallagher, 2014) investigating PSP is an important client assessment measure. To the authors' knowledge this is the *first* application of a stand-alone PSP measure to a UCC. This study demonstrated that PSP can be measured using a four-item scale, adapted from the four-item responsiveness measure of SERVQUAL (Parasuraman et al., 1988). The present study also showed that significant initial correlates of lower PSP were: clients' higher levels of mental distress when coming to counseling, waiting longer for intake, and waiting to start counseling really bothering them (wait bother). Although perhaps intuitive, these findings reinforce that it is important that UCCs systematically gather mental distress and wait-related data as part of their client assessment process (Kitzow, 2009; Valotta, 2012).

Addressing the needs of non-urgent "higher distress-wait bothered" clients may be important in order to facilitate clients presenting to their intake appointment to begin counseling. Of course, this can be challenging during UCC peak demand periods for counseling services. Alternative strategies to "hold" these clients after triage and until intake (Kitzrow, 2009) may be necessary (e.g., psychoeducational groups, drop-in counseling, self-help centers). Prior research has shown that a longer wait time after triage led to a higher no show rate at intake (DiMino & Blau, 2012). However, DiMino and Blau (2012) did not measure wait experience. This study showed that the perceived wait bother experience is more important than wait time in affecting PSP. Research is needed to investigate if wait experience is a stronger antecedent of subsequent show rate than wait time. In addition, higher PSP had a significant positive impact on the recommending the university beyond demographics, client perceptions and wait measures. This reinforces that a UCC can help with prospective student recruitment (Bishop, 2010), and be a positive source for public relations for a university (Blau et al., 2015; Brunner et al., 2014).

5. Study Limitations

A number of limitations exist. Unfortunately, data for show/no-show rate at subsequent intake could not be collected. All data are self-reported and cross-sectional, including the retrospective estimate of days waited to begin counseling after triage. It should be noted that the self-reported mean wait time of 15 days for this UCC sample was very consistent with the record-based 16-day mean wait time, using a separate prior sample at the same UCC, reported by DiMino and Blau (2012). One-item measures for mental distress level and wait bother experience do not allow for calculating reliability estimates. However, a one-factor test (Podsakoff, Mackenzie, Lee, & Podsakoff, 2003), showed that eight factors had eigenvalues over 1 and the first factor accounted for 18%

of the total variance. This suggests that method variance is not an overriding limitation for the study findings. The sample was restricted to undergraduates at a large urban, public university UCC site. Future research testing the generalizability of results, including the PSP measure, using other undergraduate samples at other UCC sites (e.g., smaller, private universities) is needed.

6. Implications for Practice and Future Research

Millennial college students often lack patience (Bland et al., 2012; Rickes, 2009). As student demand for UCC services continues to increase (Gallagher, 2014; Reetz et al., 2014), evaluating PSP and the variables affecting PSP, such as client mental distress level, wait time, and wait experience, are important to study. Moore and Owen (2014) have advocated for UCCs exploring procedures that could improve their service delivery. A UCC can establish a baseline measure of PSP and then use subsequent measures to evaluate if PSP is improving, holding steady or decreasing. Such measurement can be part of a general quality improvement assessment process (Maffini & Toth, 2017). Decreasing PSP, along with other data, may suggest that a re-assessment of UCC staffing, including additional hiring, is required (Valotta, 2012). Following the recommendation of the International Association of Counseling Services or IACS (2010, p. 20), "every effort should be made to maintain minimum staffing ratios in the range of one FTE professional staff member (excluding trainees) for every 1,000-1500 students, depending on services offered and other campus mental health agencies".

Higher PSP of a UCC can also have positive implications for the university's public relations and student recruitment (Bishop, 2010). Non-urgent undergraduate clients who waited less than 2 weeks for intake had stronger institutional commitment and a higher willingness to recommend the university to others, versus clients waiting more than 2 weeks (Blau et al., 2015). With the general availability of social media (e.g., Facebook), college students have access to increasing numbers of friends and acquaintances (Ellison, Steinfield, & Lampe, 2007). Such social media can help transfer students integrate more successfully socially into a receiving university (Nehls & Smith, 2014). Students who perceive higher PSP of a UCC may serve as "good-will ambassadors" for a university, i.e., be more likely to recommend the university to others. Research has shown that knowing someone who attended a university positively affected prospective students' perceptions of the school's atmosphere, e.g., "it feels right", "it's where I belong" (Pampaloni, 2010, p. 41). Consistently demonstrating a relationship between PSP and recruitment outcomes over time may help a UCC to leverage this relationship for more university resources.

7. Conclusion

The overall goal of this study was to explore the usefulness of a PSP measure for UCCs. The results of this study suggest a promising four-item scale for measuring PSP. In future research, wait experience data needs to be collected and evaluated against wait time for its impact on subsequent client show rate. The use of measures, such as PSP, in proactive research strategies can help to: demonstrate the positive impact of UCC on a University; justify a UCC's demand for resources; and provide fact-based recommendations for UCCs to help improve their services (Castonguay, Locke, & Hayes, 2011; Maffini & Toth, 2017).

References

Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246. https://doi.org/10.1037/0033-2909.107.2.238

Bielen, F., & Demoulin N. (2007). Waiting time influence on the satisfaction-loyalty relationship in services. *Managing Service Quality: An International Journal*, *17*, 174-193. https://doi.org/10.1108/09604520710735182

Bishop, J. (2010). The counseling center: An undervalued resource in recruitment, retention and risk management. *Journal of College Student Psychotherapy*, *24*, 248-260. https://doi.org/10.1080/87568225.2010.509219

Bland, H. W., Melton, B. F., Welle, P., & Bigham, L. (2012). Stress tolerance: New challenges for millennial college students. *College Student Journal*, *46*(3), 362-375.

Blau, G., DiMino, J., Sheridan, N., Stein, A., Casper, S., Chessler, M., & Beverly, C. (2015). Wait time for counseling affecting perceived stigma and attitude toward the university. *College Student Journal*, *49*(2), 280-290.

Brunner, J. L., Wallace, D. L., Reymann, L. S, Sellers, J.-J., & McCabe, A. G. (2014). College counseling today: Contemporary students and how counseling centers meeting their needs. *Journal of College Student Psychotherapy*, *28*, 257-324. https://doi.org/10.1080/87568225.2014.948770

Canel, C., & Anderson Fletcher, E. (2001). An analysis of service quality at a student health center. *International Journal of Health Care Quality Assurance*, *14*(6), 260-267. https://doi.org/10.1108/09526860110404220

Castonguay, L., Locke, B., & Hayes, J. (2011). The Center for Collegiate Mental Health: An example of a practice-research network in university counseling centers. *Journal of College Student Psychotherapy*, *25*(2), 105-119. https://doi.org/10.1080/87568225.2011.556929

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York, NY: Academic.

DiMino, J., & Blau, G. (2012). The relationship between wait time after triage and show rate for intake in a non-urgent student population. *Journal of College Student Psychotherapy*, *26*, 241-247. https://doi.org/10.1080/87568225.2012.685857

Elliott, K. M., & Healy, M. A. (2001). Key factors influencing student satisfaction related to recruitment and retention. *Journal of Marketing for Higher Education*, *10*(4), 1-11. https://doi.org/10.1300/J050v10n04_01

Ellison, N., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends": Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, *12*, 1143-1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

Gallagher, R. P. (2014). *National Survey of College Counseling 2014* (Monograph Series Number 9V). Published by The International Association of Counseling Services, Inc. Retrieved from http://www.collegecounseling.org/wp-content/uploads/NCCCS2014_v2.pdf

Golberstein, E., Eisenberg, D., & Gollust, S. E. (2008). Perceived stigma and mental health care seeking. *Psychiatric Services*, *59*(4), 392-399. https://doi.org/10.1176/ps.2008.59.4.392

Hunt, J. B., Watkins, D., & Eisenberg, D. (2012). How do college campuses make decisions about allocating resources for student mental health?: Findings from key participant interviews. *Journal of College Student Development*, *53*(6), 850-856. https://doi.org/10.1353/csd.2012.0083

IACS. (2010). *Accreditation for the University and College Counseling Services*. Alexandria, VA: International Association of Counseling Services, Inc.

Kitzrow, M. A. (2009). The mental health needs of today's college students: Challenges and recommendations. *NASPA Journal*, *46*(4), 646-660. https://doi.org/10.2202/1949-6605.5037

Komiya, N., Good, G., & Sherrod, N. (2000). Emotional openness as a predictor of college students' attitudes toward seeking psychological help. *Journal of Counseling Psychology*, *47*(1), 138-143. https://doi.org/10.1037/0022-0167.47.1.138

Lapan, R., Poynton, T., Marcotte, A., Marland, J., & Milam, C. (2017). College and Career Readiness Support scales. *Journal of Counseling & Development*, *95*(1), 77-86. https://doi.org/10.1002/jcad.12119

Lee, W., & Lambert, C. U. (2005). The effect of waiting time and affective reactions on customers' evaluation of service quality in a cafeteria. *Journal of Foodservice Business Research*, *8*, 19-37. https://doi.org/10.1300/J369v08n02_03

Maffini, C. S., & Toth, P. L. (2017). Quality improvement in university counseling centers. *Journal of College Student Psychotherapy*, *31*(1), 43-58. https://doi.org/10.1080/87568225.2016.1254000

Moore, J., & Owen, J. (2014). Assessing outcomes: Practical methods and evidence. *Journal of College Counseling*, *17*, 175-185. https://doi.org/10.1002/j.2161-1882.2014.00056.x

Mowbray, C., Megivern, D., Mandiberg, J., Strauss, S., Stein, C., Collins, K., … Lett, R. (2006). Campus mental health services: Recommendations for change. *American Journal of Orthopsychiatry*, *76*, 226-237. https://doi.org/10.1037/0002-9432.76.2.226

Much, K., Wagener, A. M., Breitkreutz, H. L., & Hellenbrand, M. (2014). Working with the Millennial Generation: Challenges facing 21st-Century students from the perspective of university staff. *Journal of College Counseling*, *17*(2), 37-47. https://doi.org/10.1002/j.2161-1882.2014.00046.x

Nehls, K., & Smith, B. D. (2014). The role of Facebook in fostering transfer student integration. *Journal of Student Affairs Research and Practice*, *51*(4), 392-406. https://doi.org/10.1515/jsarp-2014-0040

Nunnally, J. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw Hill.

Pampaloni, A. (2010). The influence of organizational image on college selection: What students seek in institutions of higher education. *Journal of Marketing for Higher Education*, *20*(1), 19-48. https://doi.org/10.1080/08841241003788037

Parasuraman, A., Zeithaml, V., & Berry, L. (1988). Multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*(1), 12-40.

Petrocelli, J. V. (2003). Hierarchical multiple regression in counseling: Common problems and possible remedies. *Measurement and Evaluation in Counseling and Development*, *36*, 9-22.

Podsakoff, P., Mackenzie, S., Lee, J., & Podsakoff, N. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903. https://doi.org/10.1037/0021-9010.88.5.879

Randall, E., & Bewick, B. (2016). Exploration of counsellors' perceptions of the redesigned service pathways: A qualitative study of a UK university counselling service. *British Journal of Guidance & Counseling*, *44*, 86-98. https://doi.org/10.1080/03069885.2015.1017801

Reetz, D. R., Krylowicz, B., & Mistler, B. (2014). *The Association for University and College Counseling Directors Annual Survey*. Retrieved from http://www.aucccd.org/assets/documents/2014%20aucccd%20monograph%20-%20public%20pdf.pdf

Rickes, P. (2009). Make way for Millenials!: How today's college students are shaping higher education space. *Planning for Higher Education*, *37*(2), 7-17.

Rockland-Miller, H., & Eells, G. (2006). The implementation of mental health clinical triage in university health services. *Journal of College Student Psychotherapy*, *20*(1), 39-51. https://doi.org/10.1300/J035v20n04_05

Senic, V., & Marinkovic, V. (2013). Patient care, satisfaction and service quality in health care. *International Journal of Consumer Studies*, *37*, 312-319. https://doi.org/10.1111/j.1470-6431.2012.01132.x

Shaffer, K. S., Love, M. M., Chapman, K. M., Horn, A. J., Haak, P. P., & Shen, C. Y. W. (2017). Walk-in triage systems in university counseling centers. *Journal of College Student Psychotherapy*, *31*(1), 71-89. https://doi.org/10.1080/87568225.2016.1254005

Smith, L. V., & Cokley, K. (2016). Stereotype threat vulnerability: A psychometric investigation of the Social Identities and Attitudes Scale. *Measurement and Evaluation in Counseling and Development*, *49*(2), 145-162. https://doi.org/10.1177/0748175615625752

Stevens, J. (1996). *Applied Multivariate Statistics for the Social Sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Valotta, L. (2012). Toward a more data-driven supervision of collegiate college centers. *Journal of American College Health*, *60*, 336-339. https://doi.org/10.1080/07448481.2012.663843

Winterrowd, E., Priniski, S. J., Achter, D., & Abhold, J. J. (2016). Correlates of satisfaction, intrapersonal learning, and academic outcomes at counseling centers in a university system. *College Student Journal*, *50*(2), 288-301.

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Cognitive Correlates of Japanese Language (Hiragana) Reading Abilities among School-Aged very Low Birth Weight Children

Motohiro Isaki[1,2], Tadahiro Kanazawa[1], Toshihiko Hinobayashi[1] & Hiroyuki Kitajima[3]

[1] Graduate School of Human Sciences, Osaka University, Osaka, Japan

[2] Japan Society for the Promotion of Science, Tokyo, Japan

[3] Osaka Women's and Children's Hospital, Osaka, Japan

Correspondence: Motohiro Isaki, Faculty of Health and Medical Sciences, Aichi Shukutoku University, 2-9, Katahira, Nagakute, Aichi, Japan. Tel: 81-561-62-4111. E-mail: isakimotohiro@gmail.com

Received: March 14, 2017 Accepted: March 31, 2017 Online Published: June 6, 2017

doi:10.5539/jedp.v7n2p33 URL: http://doi.org/10.5539/jedp.v7n2p33

Abstract

Previous studies have examined that the reading abilities of Very Low Birth Weight (VLBW) children are poorer than those of Normal Birth Weight (NBW) children. However, little is known about the cognitive functions that have been used to explain the reading problems in VLBW children. This study investigated that the effects of attention function on reading abilities in VLBW children. 23 VLBW children (mean age 9.1 years old) and 23 NBW children (mean age 9.2 years old) completed a reading test (containing word reading and non-word reading tasks), attention tasks, a phonological task and a naming task. The group differences were significant for the non-word reading task and attention tasks. Moreover, there were significant correlations between scores on the reading test and those on attention tasks. Multiple stepwise regression analysis suggested the reading scores were influenced by attention. These results of the present study suggest that attentional dyslexia is a characteristic of reading among VLBW children.

Keywords: very low birth weight children, preterm children, reading ability, attention

1. Introduction

Children born at Very Low Birth Weight (VLBW; <1,500g) are at risk of poor academic attainment (Johnson, Wolke, Hennessy, & Marlow, 2011; Litt et al., 2012). In a comparative survey of four countries—the United States, Canada, Germany and the Netherlands—over half of the extremely low birth weight children (children with birth weight under 1000g) needed special education services or had repeated a grade (Saigal et al., 2003). Severely poor academic attainment are associated with the risk of learning disabilities, and the incidence of learning disabilities in VLBW children is higher than that in Normal Birth Weight (NBW) children (Litt, Taylor, Klein, & Hack, 2005; Saigal, Rosenbaum, Szatmari, & Campbell, 1991). Moreover, Stanton-Chapman, Chapman, and Scott (2001) reported that the Apgar score (at 5 minutes after birth) correlated to the diagnosis of learning disabilities. The central symptom of learning disabilities is reading disorder. This study aimed to examine reading disorder among VLBW children.

1.1 Reading Disorders of VLBW Children

Many studies have found that the reading abilities of preterm/VLBW children are poorer than those of term/NBW children (Aarnoudse-Moens et al., 2009; Kovachy, Adams, Tamaresis, & Feldman, 2015). However, the degree of severity of such reading problems differs across studies. These differences are caused by factors such as the types of reading assessments implemented and age of the sample. Extremely preterm children experienced equal deficits in both word and pseudo word reading at 11 years of age (Johnson et al., 2011). At school age, very preterm children performed more poorly on complex word reading than on simple word reading (Aarnoudse-Moens, Oosterlaan, Duivenvoorden, Bernard van Goudoever, & Weisglas-Kuperus, 2011). Additionally, at 8 years of age, VLBW children had no deficit in reading accuracy, but exhibited impairments in reading speed (Guarini et al., 2010). Furthermore, while some research argues that the severity of reading disabilities reduces with age (Samuellson et al., 2006), the tendency of catch-up is limited to simple words, and no improvement was seen in complex words even with age (Aarnoudse-Moens et al., 2011).

Most studies for reading in VLBW children have conducted with alphabetic languages. There are very few studies about reading abilities of VLBW children in non-alphabetic writing cultures. This research addresses the reading characteristics of VLBW children whose language is Japanese, which does not use alphabets. The Japanese language makes use of two character forms: *kana* characters (*hiragana* and *katakana*) and *kanji*. In *kana*, each character represents a single sound (a mora, to be precise). Two-mora words are expressed by combining two characters (see Figure 1). On the other hand, *kanji* expresses morphemes as well as sounds. A *kanji* has multiple pronunciations based on how it is used within a sentence (see Figure 2). Moreover, as compared to *kana*, *kanji* has a larger number of characters, more complex character forms. While Japanese children learn all *kana* characters in the first grade, they continue to study *kanji* until the completion of middle school (the ninth grade).

1.2 Cognitive Function That Affect the Reading of VLBW Children

Figure 1. An example of Japanese *hiragana*

Note. A word spoken [uma] phonetically is written like Figure 1. The sound of the first letter of Figure 1a is [u], the sound of the last letter is [ma].

Figure 2. An example of Japanese *kanji*

Note. A kanji like Figure 2 means horse. However, its pronunciation varies from [uma], [ba] to [ma], depending on the sentence.

The purpose of this study was to determine the core cognitive factors that influence the reading deficits among VLBW children. The double deficit hypothesis (Wolf & Bowers, 1999) is a well-known explanation of the causes of dyslexia. The double deficit hypothesis suggests that poor readers show deficits in either phonological processing or naming speed, or in both. However, in recent years, there have been studies postulating that attention also influences reading abilities. Some studies indicate that children who experience reading difficulties have problems with visuospatial attention skills (Bosse, Tainturier, & Valdois, 2007; Franceschini, Gori, Ruffino, Pedrolli, & Facoetti, 2012). Given these findings, some researchers discuss the existence of attentional dyslexia, the causes of which differ from the dyslexic symptoms indicated by the double deficit hypothesis (Elliott & Grigorenko, 2014). In addition, VLBW children have been reported to exhibit attention deficits (Anderson et al., 2011; Johnson & Wolke, 2013; Mulder, Pitchford, Hagger, & Marlow, 2009). Therefore, in studying the cause of reading impairments in VLBW children, attentional dyslexia may be a highly relevant concept.

Although, little is known about the cognitive functions that have been used to explain the reading problems in VLBW children, Wocadlo and Rieger (2007) examined the reading abilities of very preterm children using the double deficit hypothesis. They indicated that very preterm children who had difficulties with phonological processing or naming speed scored lower on reading accuracy than did very preterm children who did not have such difficulties. Further, those who had problems with both phonological processing and naming speed had even poorer reading scores. Thus, they revealed that the double deficit hypothesis was also applicable to the reading abilities of very preterm children. However, as their research did not make comparisons with full term and/or normal birth weight children, it has not been clarified whether VLBW children's scores on reading accuracy, phonological processing, and naming speed were lower than those of NBW children. Moreover, they suggested that attention is likely associated to their reading performance. Nevertheless, they did not conduct a separate test of attention.

Very few studies have compared VLBW children's phonological processing and naming speed with NBW children. With regard to phonological processing, VLBW children's scores have been shown to be lower than those of NBW children (Guarini et al., 2010; Johnson et al., 2011; T. Luu, Vohr, Allan, Schneider, & Ment, 2011; Mullen et al., 2011). On the other hand, there have been no differences shown in naming speed between VLBW

(or preterm) children and NBW (or term) children (Luu et al., 2009; Mullen et al., 2011; Myers et al., 2010; Saavalainen et al., 2006).

Studies on VLBW children's brain functions suggest that attention influences reading. A study using magnetoencephalography (MEG) showed differences in the areas of brain activity between term children and preterm children. Preterm children particularly those with above average reading scores, demonstrated an over activation in the Broca's area, which is linked to language skills, and the prefrontal area, which is linked to attention. Further, poor activation of the prefrontal area was observed in preterm children who had reading difficulties (Frye et al., 2009). Nevertheless, only a few studies have clarified the relation between VLBW children's reading and attention using psychological measurement methods. While Anderson et al. (2011) suggest that the function of attention is related to VLBW children's reading, they did not conduct a reading test. According to Jaekel, Wolke, and Bartmann (2013), behavioral observation of inattention is related to academic outcomes. Yet, this study did not utilize any neuropsychological test of attention. In other words, very few studies on the relation between reading abilities and attentional functions of VLBW children used objective methods to investigate either of these abilities. Therefore, the present study aims to explore VLBW children's reading, and cognitive functions, which may have an influence on reading. It will explore the hypotheses that there are no group differences in the phonological processing and naming speed task scores, but that the scores of attention would be lower in VLBW children than that in NBW children; and phonological processing, naming speed and attention will influence reading abilities.

2. Method

2.1 Participants

Participants were 23 VLBW children (<1,500g birth weight, without intellectual disabilities or any apparent physical, visual, or hearing disabilities) and 23 NBW children (control group). All VLBW children were born in Hospital A (17 male, 6 female; mean age 9.1 years, mean birth weight 822.7g, mean gestation period 27.2 weeks). Table 1 indicates the perinatal characteristics of the VLBW group. The NBW group included 23 children who attended Primary School B (16 male, 7 female; mean age 9.2 years). All NBW children were healthy and with no history of perinatal problems and developmental disabilities. To make the groups as comparable as possible, we eliminated the participants with intellectual disabilities, cerebral palsy, or visual disabilities. Independent sample *t* test revealed that VLBW group and NBW group sample did not differ significantly on their ages.

Table 1. Medical characteristics of the VLBW group

	Mean	SD
Birth weight (g)	822.7	263.1
Gestational age (weeks)	27.2	3.0
Head circumference at birth (cm)	23.9	2.5
Maternal age	30.7	4.0
Apgar 1 minute	4.3	1.9
Apgar 5 minute	7.2	1.8
Time on respirator (days)	36.7	28.9
Time on Oxygen (days)	38.4	42.9
		number
intraventricular hemorrhage [1)]		4
periventricular leukomalacia		1
bronchopulmonary dysplasia		15
retinopathy of prematurity		19

Note 1. IVH grade is **[**.

In order to ensure the similarity of two groups' intelligence, we conducted the "block design" and "words" subtests of the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-Ⅳ). This serves as a screening method to estimate group differences in intellectual abilities, and has also been used in a previous study targeting VLBW children (Kulseng et al., 2006). The groups did not show any differences at 5% level. Thus, the two samples did not differ in terms of intellectual abilities.

2.2 Ethical Considerations and Recruitment Procedure

This research obtained approval from the Osaka University Graduate School of Human Sciences Behavioral Sciences Research Ethics Committee (No. 24066). Additionally, for the sample of VLBW children, we received ethical approval from the Ethics Committee of Hospital A. An in-clinic psychologist explained this study during regular follow-ups at Hospital A, and we conducted assessments on children who consented to participate in this study. Furthermore, we also obtained approval to implement the study from the educational committee of Primary School B's municipality. The principal explained the study in a parent-teacher meeting, and consent forms were distributed. We conducted assessments on children who gave consent.

2.3 Neuropsychological Measurement Items

2.3.1 Measurement of Reading Abilities

Reading abilities were assessed using the "reading test" (Inagaki, 2010). In this task, the child needs to read aloud thirty words and non-words each, consisting of four *hiragana* characters. We recorded the time required for oral reading to indicate fluency, and the number of reading mistakes to assess accuracy of decoding.

2.3.2 Measurement of Attention

Visual search tasks are often utilized as measurements of attention in reading research. This study employed the Sky Search and Map Mission tasks from the Test of Everyday Attention for Children (TEA-Ch: Manly, Robertson, Anderson, & Nimmo-Smith, 1999). TEA-Ch has not been standardized in Japan, so we translated as we went. In both tasks, the child is asked to search for target visual stimuli. While the Sky Search task has no set time limitations, the Map Mission task must be completed within one minute.

2.3.3 Measurement of Phonological Processing Abilities

There are no standardized tests to assess phonological processing in Japan. Phoneme deletion task and word reversal task are the recommended methods to assess phonological processing (Hara, 2003). The phoneme deletion task would be easy given the age of the participants in this study, and is thus highly likely to demonstrate a ceiling effect. So this study utilized the word reversal task. For this task, we used 15 words, containing two- to four-mora words. For example, [tanuki] (a word meaning raccoon) is a three-mora word, and in reverse, [kinuta] is a correct response. The word [zenbu] (meaning all) is a two-syllable word, but it is also a three-mora word. Therefore, in reverse, [bunze] is correct. Each stimulus word was presented vocally, and the children were asked to answer orally. We measured the number of mistakes.

2.3.4 Measurement of Naming Speed Ability

There is no standardized testing for naming speed ability in Japan. We utilised the Rapid Automated Naming (RAN) task by Kaneko, Uno, Haruhara (2004). We prepared three A4-size papers with ten line drawings and ten numbers placed randomly (each with different drawings and numbers). We measured the total time taken to name the contents on the three sheets. We also measured the number of mistakes. However, we did not record any mistakes in all the participants of this experiment.

2.4 Statistical Analysis

We performed independent *t*-tests to analyze group differences in the neuropsychological measurement tasks. To investigate the impact of cognitive functions that influence on word reading, we ran a multiple regression analysis using fluency scores of reading as the objective variable and the scores on Sky Search task, Map Mission task, word reversal task and RAN task as the explanatory variable. To assess model fit, we applied a stepwise procedure. We used Akaike's Information Criterion (AIC) to indicate good fit model. All statistical analyses were conducted using R3.1.2.

3. Results

3.1 Results of Neuropsychological Measurement Items

Table 2 indicates the measurement outcomes of the VLBW and NBW groups.

Table 2. Scores for neuropsychological tests in VLBW and NBW children

Neuropsychological tests		VLBW (n=23) (m=17, f=6)		NBW (n=23) (m=16, f=7)		
		Mean	SD	Mean	SD	*p*
Reading accuracy	Words (errors)	0.3	0.5	0.2	0.4	0.24
	Nonwords (errors)	2.3	2.3	1.1	1.1	0.02
Reading fluency	Words (sec)	30.9	10.2	33.5	13.3	0.22
	Nonwords (sec)	64.9	14.1	57.3	16.6	0.04
Phonological processing naming speed attention	Word reversal (errors)	2.3	2.2	2.6	2.6	0.38
	RAN (sec)	44.9	10.7	41.5	8.2	0.15
	Sky search (duration/correct responses)	6.2	2.2	5.5	1.1	0.09
	Map mission (correct responses)	23.6	5.0	27.8	7.3	0.02

3.1.1 Reading Test

There were no group differences in the scores of word reading task (accuracy: $t(44)=0.70$, $p=0.24$; fluency: $t(44)=0.76$, $p=0.22$). On the other hand, the group differences were significant for non-word reading task (accuracy: $t(44)=2.19$, $p<0.05$; fluency: $t(44)=1.68$, $p<0.05$).

3.1.2 Measurement Outcomes of Cognitive Functions

There were no group differences in the scores of the word reversal task and RAN task, which are measured in the double deficit hypothesis of reading. The group differences were significant for Sky Search task and Map Mission task. In other words, the VLBW and NBW groups showed no significant differences in phonological processing task and naming speed task, but the VLBW group scored significantly lower than the NBW group in the attention tasks.

3.2 Cognitive Functions That Influence Reading Abilities

First, we explored the association between the scores of neuropsychological tasks. There were significant correlations between scores on reading test and those on double deficit (word reversal task and RAN task). With regard to selective attention measurements, there were significant correlations between scores on reading test and those on attention tasks (refer to Table 3).

Table 3. Correlations among neuropsychological tests

Neuropsychological tests	1	2	3	4	5
Reading fluency					
1. Words					
2. Nonwords	0.75***				
Double deficit for reading					
3. Words reversal (phonological processing)	0.57***	0.53***			
4. RAN task (naming speed)	0.58***	0.57***	0.49***		
Attention					
5. Sky serach	0.36*	0.27†	-0.01	0.18	
6. Map mission	-0.15	-0.31*	0.10	-0.24	-0.30*

Note. †<0.1, *p<0.05, **p<0.01, ***p<0.001.

Next, multiple stepwise regression analysis revealed the cognitive factors that influence reading scores. Table 4 indicates the outcomes. The word reversal, RAN, and Sky Search influenced word reading. The coefficient of determination for this model was R^2=0.51 at 0.1% level of significance [F(3, 42)=16.53, p<0.0001]. Similarly, the word reversal, RAN, and Map Mission influenced non-word reading. The coefficient of determination for this model was R^2=0.44 at 0.1% level of significance [F(3, 42)=12.83, p<0.0001]. Both these models suggest that the reading scores were influenced not only by phonological awareness and naming speed (the double deficit hypothesis) but also by attentional functions.

Table 4. Neuropsychological abilities independently associated with reading from multivariable stepwise regression analysis

Variables	R^2	β	95%*CI*	*p*
Words	0.51			
Word reversal (phonological processing)		0.43	[0.18-0.69]	0.001
RAN task (naming speed)		0.32	[0.07-0.56]	0.01
Sky search (attention)		0.31	[0.09-0.52]	0.006
Nonwords	0.44			
Word reversal (phonological processing)		0.41	[0.15-0.68]	0.003
RAN task (naming speed)		0.30	[0.03-0.57]	0.03
Map mission (attention)		-0.28	[-0.52--0.04]	0.02

Note. The β value of Map mission is minus quantity because higher scores for Map mission task are indicative of better performance but lower scores for other tasks are indicative of better.

4. Discussion

This study rectified many of the limitations of previous research about reading abilities of VLBW children. By conducting many neuropsychological tests, we were able to provide new knowledge about reading abilities in VLBW children.

4.1 Reading Abilities of VLBW Children

VLBW children performed more poorly than NBW children on accuracy and fluency for non-word reading. In contrast, VLBW children did not score differently from NBW children for word reading. According to Shaywitz (2003), non-word reading is probably the best measurement of pure decoding abilities. These findings suggest that VLBW children exhibit deficits in decoding.

Guarini et al. (2010) did not show significant differences of accuracy for non-word reading between VLBW and NBW group. The sample used in Guarini et al.'s (2010) study had a mean birth weight of 1243g and a mean gestational period of 30.4 weeks. Whereas, the sample used in this study had a mean birth weight of 822.7g and a mean gestational period of 27.2 weeks. In other words, the children in this study were born smaller and earlier, which may have contributed to further deficits in accuracy. For example, the prevalence rate of reading disorders in late preterm (gestational period over 34 weeks but under 37 weeks) children was not significantly different from that in term children (Harris et al., 2013). Additionally, a meta-analysis on the reading abilities of preterm children with gestational age under 32 weeks found that the smaller the mean gestational age of the sampled children in the data, the lower their reading abilities were (Kovachy et al., 2015). Moreover, a meta-analysis of children with gestational age under 33 weeks and birth weight under 1500g revealed a correlation between gestational age and birth weight, and reading scores (Aarnoudse-Moens, Weisglas-Kuperus, van Goudoever, & Oosterlaan, 2009). In other words, these findings suggest that the smaller the gestational age and lighter the birth weight, the higher is the risk of reading disorders. We consider that the differences in outcomes of our study and those of Guarini et al. (2010) are due to the differences in the birth weight and gestational age of the VLBW children sampled.

Next, we consider the reasons for the different outcomes of fluency for word reading in this study and that of Guarini et al. (2010). The Italian version of the reading test, which Guarini et al. used for their study, included 112 words and 48 non-words, while the Japanese version of the reading test used in this research had 30 words

each for words and non-words. Such differences in the number of items in the test could have influenced the outcome of fluency. In other words, the fluency of VLBW children could deteriorate as the number of items increase. This is due to the fact that VLBW children have difficulties maintaining attention, as indicated by preceding studies (Anderson et al., 2011; Mulder et al., 2009). The difficulties of fluency in VLBW children may not be apparent while they read fewer words, as observed in the present study; but when the number of words increases, their fluency deteriorates as they are unable to maintain the attention required for reading.

4.2 Cognitive Functions That Influence VLBW Children's Reading

According to double deficit hypothesis (Wolf & Bowers, 1999), phonological processing and naming speed relate to reading ability. Targeting 8 year-old very preterm children, Wocadlo and Rieger (2007) clarified that the double deficit hypothesis also applies to preterm children. However, their study did not compare with the term groups, and did not clarify if preterm children's phonological processing and naming speed are impaired. Therefore, the present study conducted comparisons between NBW and VLBW children, and examined whether there are group differences in phonological processing and naming speed abilities. Additionally, this study also investigated the relationship between attention and reading.

This study showed that the scores of phonological processing and naming speed in VLBW children were not different from those of NBW children. With regards to naming speed, group differences have not been reported in preceding studies either (Luu et al., 2009; Mullen et al., 2011; Saavalainen et al., 2006), thus matching the outcome of the present study. Confirming that naming speed abilities are less likely to be impaired in school-aged VLBW children.

With regard to phonological processing abilities, previous studies (Guarini et al., 2010; Johnson et al., 2011; Luu, Vohr, Allan, Schneider, & Ment, 2011; Mullen et al., 2011) recognize group differences, which was not reflected in the present findings. This inconsistency between previous studies and the present study may relate the character system of language. While the previous studies targeted children in alphabetic writing cultures, the present study targeted children in the Japanese language (non-alphabetic) culture. In recent years, some researchers have argued that phonological processing may not have as much influence on reading abilities in the Japanese language (Welty, Menn, & Oishi, 2014). On the other hand, some studies have reported phonological processing scores to be correlated to the fluency and accuracy scores for reading (Seki, Kassai, Uchiyama, & Koeda, 2008). This demonstrates a need for further research regarding phonological processing in the Japanese language.

Next, this research hypothesized that attention relate to VLBW children's reading, and we were able to verify this hypothesis. The link between preterm children's reading abilities and attention has been indicated in studies using brain imaging (Frye et al., 2009, 2010), but there are few studies that indicate this relationship using a psychological measurement. The present study supports findings of a study conducted by Frye et al. (2009, 2010) using psychological methods. Our results showed that the score of selective attention in neuropsychological test influenced reading abilities. Valdois et al. (2011) reported that there are children who have reading disorders due to attention problems despite not having a phonological processing problem (Valdois et al., 2011). In other words, the results of the present study suggest that attentional dyslexia (Elliott & Grigorenko, 2014) might be a characteristic of reading among VLBW children.

4.3 The Limitations of the Present Study and Future Prospects

The present study investigated the reading abilities of Japanese VLBW children, but the reading tests were limited to the reading of *hiragana*, and did not include *kanji*. As compared to *hiragana*, *kanji* has complex character forms. As VLBW children scored lower on complex word reading as compared to simple word reading in alphabetical writing forms (Aarnoudse-Moens et al., 2011), they may score lower on *kanji* reading than on *hiragana* reading. Furthermore, *kanji* uses letters from the Chinese language. Interestingly, the brain areas activated during reading are different for readers of an alphabetic language and the Chinese language, with Chinese language readers demonstrating more brain activity in the left middle frontal gyrus. This area is related to visual attention, and suggests that more visual attention is required to read the complex letters of the Chinese language than for reading an alphabetical language (Siok, Perfetti, Jin, & Tan, 2004). In other words, it is possible that reading *kanji* could require more visual attention than reading *hiragana* does. The need to clarify the relationship between *kanji* reading and attention among VLBW children is a central theme in future research.

Acknowledgements

This study was supported by the Japan Society for the Promotion of Science (JPSS KAKENHI Grant Number JP 26•1571). We would like to thank pediatrician Shinya Hirano and psychologists Etsuyo Yamamoto and Machiko Tajima in Osaka Women's and Children's Hospital, in addition to the families who participated in this study.

References

Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics*, *124*(2), 717-728. https://doi.org/10.1542/peds.2008-2816

Aarnoudse-Moens, C., Oosterlaan, J., Duivenvoorden, H., Bernard van Goudoever, J., & Weisglas-Kuperus, N. (2011). Development of preschool and academic skills in children born very preterm. *The Journal of Pediatrics*, *158*(1), 51-56. https://doi.org/10.1016/j.jpeds.2010.06.052

Anderson, P. J., De Luca, C. R., Hutchinson, E., Spencer-Smith, M. M., Roberts, G., & Doyle, L. W. (2011). Attention problems in a representative sample of extremely preterm/extremely low birth weight children. *Developmental Neuropsychology*, *36*(1), 57-73. https://doi.org/10.1080/87565641.2011.540538

Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition*, *104*(2), 198-230. https://doi.org/10.1016/j.cognition.2006.05.009

Elliott, J. G., & Grigorenko, E. L. (2014). *The dyslexia debate*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139017824

Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Current Biology*, *22*(9), 814-819. https://doi.org/10.1016/j.cub.2012.03.013

Frye, R. E., Hasan, K., Malmberg, B., Desouza, L., Swank, P., Smith, K., & Landry, S. (2010). Superior longitudinal fasciculus and cognitive dysfunction in adolescents born preterm and at term. *Developmental Medicine and Child Neurology*, *52*(8), 760-766. https://doi.org/10.1111/j.1469-8749.2010.03633.x

Frye, R. E., Malmberg, B., deSouza, L., Swank, P., Smith, K., & Landry, S. (2009). Increased prefrontal activation in adolescents born prematurely at high risk during a reading task. *Brain Research*, *1303*, 111-119. https://doi.org/10.1016/j.brainres.2009.09.091

Guarini, A., Sansavini, A., Fabbri, C., Savini, S., Alessandroni, R., Faldella, G., & Karmiloff-Smith, A. (2010). Long-term effects of preterm birth on language and literacy at eight years. *Journal of Child Language*, *37*, 865-885. https://doi.org/10.1017/S0305000909990109

Harris, M. N., Voigt, R. G., Barbaresi, W. J., Voge, G. A., Killian, J. M., Weaver, A. L., … Katusic, S. K. (2013). ADHD and Learning Disabilities in Former Late Preterm Infants: A Population-Based Birth Cohort. *Pediatrics*, *132*(3), 630-636. https://doi.org/10.1542/peds.2012-3588

Jaekel, J., Wolke, D., & Bartmann, P. (2013). Poor attention rather than hyperactivity/impulsivity predicts academic achievement in very preterm and full-term adolescents. *Psychological Medicine*, *43*, 183-196. https://doi.org/10.1017/S0033291712001031

Johnson, S., & Wolke, D. (2013). Behavioural outcomes and psychopathology during adolescence. *Early Human Development*, *89*(4), 199-207. https://doi.org/10.1016/j.earlhumdev.2013.01.014

Johnson, S., Wolke, D., Hennessy, E., & Marlow, N. (2011). Educational outcomes in extremely preterm children: Neuropsychological correlates and predictors of attainment. *Developmental Neuropsychology*, *36*(1), 74-95. https://doi.org/10.1080/87565641.2011.540541

Kovachy, V. N., Adams, J. N., Tamaresis, J. S., & Feldman, H. M. (2015). Reading abilities in school-aged preterm children: A review and meta-analysis. *Developmental Medicine & Child Neurology*, *57*(5), 410-419. https://doi.org/10.1111/dmcn.12652

Kusleng, S., Jennekens-Schinkel, A., Naess, P., Romundstad, P., Indredavik, M., Vik, T., & Brubakk, A.-M. (2006). Very-low-birthweight and term small-for-gestational-age adolescents: Attention revisited. *Acta Paediatrica*, *95*(2), 224-230.

Litt, J. S., Taylor, H. G., Margevicius, S., Schluchter, M., Andreias, L., & Hack, M. (2012). Academic achievement of adolescents born with extremely low birth weight. *Acta Paediatrica (Oslo, Norway: 1992)*, *101*, 1240-1245. https://doi.org/10.1111/j.1651-2227.2012.02790.x

Litt, J., Taylor, H. G., Klein, N., & Hack, M. (2005). Learning Disabilities in Children with Very Low Birthweight: And Educational Interventions. *Journal of Learning Disabilities*, *38*, 130-141. https://doi.org/10.1177/00222194050380020301

Luu, T. M., Ment, L. R., Schneider, K. C., Katz, K. H., Allan, W. C., & Vohr, B. R. (2009). Lasting effects of preterm birth and neonatal brain hemorrhage at 12 years of age. *Pediatrics*, *123*, 1037-1044. https://doi.org/10.1542/peds.2008-1162

Luu, T. M., Vohr, B. R., Allan, W., Schneider, K. C., & Ment, L. R. (2011). Evidence for catch-up in cognition and receptive vocabulary among adolescents born very preterm. *Pediatrics*, *128*(2), 313-322. https://doi.org/10.1542/peds.2010-2655

Manly, T., Robertson, I. H., Anderson, V., & Nimmo-Smith, I. (1999). *The Test of Everyday Attention for Children*. London: Pearson.

Mulder, H., Pitchford, N. J., Hagger, M. S., & Marlow, N. (2009). Development of executive function and attention in preterm children: A systematic review. *Developmental Neuropsychology*, *34*(4), 393-421. https://doi.org/10.1080/87565640902964524

Mullen, K. M., Vohr, B. R., Katz, K. H., Schneider, K. C., Lacadie, C., Hampson, M., … Ment, L. R. (2011). Preterm birth results in alterations in neural connectivity at age 16 years. *NeuroImage*, *54*(4), 2563-2570. https://doi.org/10.1016/j.neuroimage.2010.11.019

Myers, E. H., Hampson, M., Vohr, B., Lacadie, C., Frost, S. J., Pugh, K. R., … Ment, L. R. (2010). Functional connectivity to a right hemisphere language center in prematurely born adolescents. *NeuroImage*, *51*(4), 1445-1452. https://doi.org/10.1016/j.neuroimage.2010.03.049

Saavalainen, P., Luoma, L., Bowler, D., Timonen, T., Määttä, S., Laukkanen, E., & Herrgård, E. (2006). Naming skills of children born preterm in comparison with their term peers at the ages of 9 and 16 years. *Developmental Medicine & Child Neurology*, *48*, 28-32. https://doi.org/10.1017/s0012162206000077

Saigal, S., den Ouden, L., Wolke, D., Hoult, L., Paneth, N., Streiner, D. L., … Pinto-Martin, J. (2003). School-age outcomes in children who were extremely low birth weight from four international population-based cohorts. *Pediatrics*, *112*, 943-950. https://doi.org/10.1542/peds.112.4.943

Saigal, S., Rosenbaum, P., Szatmari, P., & Campbell, D. (1991). Learning disabilities and school problems in a regional cohort of extremely low birth weight (less than 1000 G) children: A comparison with term controls. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *12*, 294-300. https://doi.org/10.1097/00004703-199110000-00002

Samuelsson, S., Finnström, O., Flodmark, O., Gäddlin, P. O., Leijon, I., & Wadsby, M. (2006). A longitudinal study of reading skills among very-low-birthweight children: Is there a catch-up? *Journal of Pediatric Psychology*, *31*, 967-977. https://doi.org/10.1093/jpepsy/jsj108

Sansavini, A., Guarini, A., Savini, S., Broccoli, S., Justice, L., Alessandroni, R., & Faldella, G. (2011). Longitudinal trajectories of gestural and linguistic abilities in very preterm infants in the second year of life. *Neuropsychologia*, *49*, 3677-3688. https://doi.org/10.1016/j.neuropsychologia.2011.09.023

Seki, A., Kassai, K., Uchiyama, H., & Koeda, T. (2008). Reading ability and phonological awareness in Japanese children with dyslexia. *Brain and Development*, *30*, 179-188. https://doi.org/10.1016/j.braindev.2007.07.006

Shaywitz, S. (2003). *Overcoming dyslexia. A new and complete science-based program for reading problems at any level*. New York: Random House.

Stanton-Chapman, T. L., Chapman, D. A., & Scott, K. G. (2001). Identification of Early Risk Factors for Learning Disabilities. *Journal of Early Intervention*, *24*(3), 193-206. https://doi.org/10.1177/105381510102400305 01

Valdois, S., Bidet-Ildei, C., Lassus-Sangosse, D., Reilhac, C., N'guyen-Morel, M. A., Guinet, E., & Orliaguet, J. P. (2011). A visual processing but no phonological disorder in a child with mixed dyslexia. *Cortex*, *47*(10), 1197-1218. https://doi.org/10.1016/j.cortex.2011.05.011

Welty, Y. T., Menn, L., & Oishi, N. (2014). Developmental Reading Disorders in Japan—Prevalence, Profiles, and Possible Mechanisms. *Topics in Language Disorders*, *34*(2), 121-132. https://doi.org/10.1097/TLD.0000000000000014

Wocadlo, C., & Rieger, I. (2007). Phonology, rapid naming and academic achievement in very preterm children at eight years of age. *Early Human Development*, *83*, 367-377. https://doi.org/10.1016/j.earlhumdev.2006.08.001

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*, 415-438. https://doi.org/10.1037/0022-0663.91.3.415

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Does the Number of Available Strategies Change How Children Perform Cognitive Tasks? Insights from Arithmetic

Patrick Lemaire[1], Koen Luwel[2] & Fleur Brun[1]

[1] Aix-Marseille University, CNRS, LPC, Marseille, France

[2] KU Leuven-Campus Brussels, Brussels, Belgium

Correspondence: Patrick Lemaire, Aix-Marseille University & CNRS, 3, Place Victor Hugo, Case D, 13331, Marseille, France. E-mail: patricklemaire13@gmail.com

Received: May 14, 2017 Accepted: June 8, 2017 Online Published: June 28, 2017

doi:10.5539/jedp.v7n2p43 URL: http://doi.org/10.5539/jedp.v7n2p43

Abstract

Fifth and seventh-graders accomplished computational estimation tasks in conditions where only one versus two strategies were available. Children were told which strategy to execute on each problem. Results showed that both groups of children were faster under one-strategy condition than under two-strategy condition and that age-related differences in performance were larger under two-strategy condition. Also, differences in strategy performance tended to vary as a function of the number of strategies, and this strategy difference was largest in younger children. These findings have implications to further our understanding of strategy execution in arithmetic and in other cognitive domains, as well as of age-related differences in children's performance during cognitive development.

Keywords: arithmetic, computational estimation, elementary-school children, strategies

1. Introduction

In many cognitive domains, children use a variety of strategies (Siegler, 2007). A strategy is "a procedure or a set of procedures to achieve a higher level goal" (Lemaire & Reder, 1999, p. 365). Children's cognitive performance as well as age-related differences in this performance crucially depend on which strategies children use, how often they use each available strategy, and how efficient they are at selecting and executing strategies (see Siegler, 1996, for an overview). In this study, we focused on strategy execution. Theoretical and empirical works have shown that relative strategy difficulty depends on the number and type of procedures included within each strategy: Strategies including more and/or harder procedures yield longer latencies and higher error rates. Previous findings also revealed that age-related differences were largest when children use harder strategies than easier strategies and that relative strategy performance is influenced by situation, person, and problem characteristics (e.g., Siegler, 2007). Unknown is whether age-related differences in children's cognitive performance are influenced by the number of available strategies. The present experiment addressed this issue in the context of arithmetic problem solving.

When children are given either simple (e.g., 3x4; 5+7) or more complex (e.g., 27-9; 38+46) arithmetic problems, they use several strategies. For example, in computational estimation tasks, investigated here, when children are asked to provide approximate sums to problems like 52+47, they use several rounding strategies, like rounding both operands down or up to their closest decades (e.g., 50+40, 60+50), or like rounding one operand down and the other up to its closest decade (e.g., 50+50). These strategies are known and spontaneously used by children as young as seven years old, and different strategies yield different levels of speed and accuracy (e.g., LeFevre, Greenham, & Waheed, 1993). For example, rounding both operands down is faster and more accurate (i.e., estimated and correct sums are closer) on small-unit problems (i.e., problems with sum of unit digits <10, like 41+62) and rounding both operands up is most efficient on large-unit problems (i.e., problems with sum of unit digits >10, like 37+49). Relative strategy performance has been found to vary not only with problem features, but also with children's characteristics (e.g., age, arithmetic skills) and with situational demands like the need to respond quickly and/or accurately (e.g., Dowker, 1997; Dowker, Flod, Griffiths, Harriss, & Hook, 1996; LeFevre et al., 1993; Lemaire & Lecacheur, 2002; Reys, Rybolt, Bestgen, & Wyatt, 1982; Sowder & Markovits, 1990). For example, Lemaire and Brun (2014; see also Lemaire & Lecacheur, 2011) found that when children

accomplished computational estimation tasks, young children executed the rounding-down strategy more quickly on small-unit problems (i.e., doing 30+40 to estimate 32+46) than on large-unit problems (i.e., doing 20+30 to estimate 24+39), and that this Strategy x Problem interaction was stronger in third than in fifth graders. Similarly, Lemaire and Brun (2014) found that age-related differences in latencies to execute the rounding-down and the rounding-up strategies were larger under short response-stimulus interval condition (i.e., when the next problem was displayed 900 ms after participants' answer) than under longer response-stimulus interval condition (i.e., the next problem was displayed 1900 ms after participants' answer). In other words, both age-related differences and relative strategy performance vary as a function of problem type and situational characteristics. Unknown however is whether age- and strategy-related differences are influenced by the number of strategies that children use.

The goal of the present experiment was to determine whether effects of children's age and of strategies on cognitive performance are influenced by how many strategies are used. This is important as it would have a number of significant implications to further our understanding of both strategic behaviors and age-related differences in performance during children's cognitive development. For example, one of the important theoretical implications concerns determiners of relative strategy performance. Computational models of strategies (Lovett and Anderson (1996)'s ACT-R model; Lovett and Schunn (1999)'s RCCL model; Payne, Bettman, and Johnson (1993)'s adaptive decision maker model; Rieskamp and Otto (2006)'s SSL model; and Siegler and Arraya (2005)'s SCADS* model; see overview by Marewski and Link, 2014) share the core assumption that relative strategy performance depends on the number and types of mental procedures involved in each strategy. That is, individuals are faster with strategies that involve fewer and/or easier procedures. Finding that relative strategy performance differs in conditions where children use one versus two strategies would suggest that strategy execution is also influenced by some contextual factors, like the number of strategies brought to the task, above and beyond the component processes of each strategy.

From a developmental perspective, in most cognitive domains, age-related differences in children's cognitive performance are larger for harder than for easier strategies. For example, in arithmetic, many researchers found that age-related differences in performance are larger when children use harder, counting strategies relative to when they use an easier, direct retrieval strategy to solve arithmetic problems like 8+4 (see Cohen-Kadosh & Dowker, 2015; Geary, 1994, for overviews). Finding that these age-related differences for easier and harder strategies depend on the number of strategies that are used would imply that age-related improvements in children's performance are not only the result of increased efficacy at executing strategies but also at managing multiple strategies.

In the present study, fifth and seventh graders were asked to find sum estimates to two-digit addition problems. For each problem, a cue indicated which strategy children had to execute, which means that children did not have to select a strategy themselves. We controlled strategy selection to assess the role of the number of strategies on children's performance. By doing so, we avoided that strategy execution would be contaminated by differences in the frequency with which children would use each of the available strategies as well as differences in the type of problems on which the respective strategies would be applied. Children were tested under both a two-strategy and a one-strategy condition. In the two-strategy condition, children solved a set of 32 problems and were asked to execute either a mixed rounding-down or a mixed rounding-up strategy on each problem. In the Mixed Rounding-Down strategy (MRD), children rounded the first operand down and the second operand up to the closest decades (e.g., doing 40+70 to estimate 43+68). In the Mixed Rounding-Up strategy (MRU), children rounded the first operand up and the second operand down to the closest decades (e.g., doing 50+60 to estimate 43+68). In the one-strategy condition, children were asked to solve a first set of 16 problems and were cued with the same strategy on all these 16 problems. Children were also asked to solve another set of 16 problems and were cued with the other strategy on all these 16 problems. MRD and MRU were tested because previous works showed that both fifth and seventh graders know and spontaneously use these two mixed-rounding strategies (LeFevre et al., 1993), and because previous works on relative strategy execution found that MRD is easier than MRU and yields better performance (e.g., Hinault, Lemaire, & Phillips, 2016; Lemaire & Brun, 2014; Lemaire & Hinault, 2014; Uittenhove & Lemaire, 2012).

As previous works found that the number of strategies individuals use to accomplish a given task is related to their executive functions (Ardiale, Hodzik, & Lemaire, 2012; Hodzik & Lemaire, 2011), we hypothesized that the number of available strategies would influence children's performance and age-related differences in this performance. We therefore tested the following predictions. First, children were expected to be slower and less accurate in the two- than in the one-strategy condition. This could happen if children have to use more executive control resources in the two-strategy condition relative to the one-strategy condition to maintain both strategies

activated in working memory and to switch between strategies across problems. Second, if the number of strategies affects strategy execution, relative strategy performance should differ in the one- and two-strategy conditions, such that speed and accuracy should decrease more strongly for the harder (MRU) strategy than for the easier (MRD) strategy from the one- to the two-strategy condition. Again, this should occur if managing two strategies requires more resources than managing only one strategy and because children need more resources to execute the harder than the easier strategy. Third, larger age-related differences were expected in the two-strategy condition relative to the one-strategy condition. This would happen if the two-strategy condition requires more processing resources that are known to increase with children's age (Lemaire & Lecacheur, 2011). Finally, a Group x Number of strategies x (MRD/MRU) Strategy interaction was expected if, relative to older children, younger children are more influenced by the number of strategies while executing the harder strategy than when using the easier strategy.

2. Method

2.1 Participants

Seventy-two children were tested: 36 fifth graders (17 girls; mean age=130 months; range=123-144) who came from one elementary school and 36 seventh graders (13 girls; mean age=153 months; range=145-165) who were drawn from two secondary schools located in Flanders, the Dutch-speaking part of Belgium. Children's parents provided written informed consent, and children were told that they could quit the experiment at any moment.

2.2 Stimuli

Two sets of 16 two-digit addition problems (e.g., 24+39) each were selected for the two-strategy condition. All problems included one operand with its unit digit smaller than five and the other operand with its unit digit larger than five. In each set of 16 problems, the unit digit of the first operand was smaller than five and the unit digit of second operand was larger than five on half the problems (and the reverse for the other problems). Both sets of 16 problems were matched on the size of correct sums and on percent deviations between estimates and correct sums.

Moreover, following previous findings in arithmetic (see Campbell, 2005; Cohen Kadosh & Dowker, 2015, for overviews), the following factors were controlled: (a) no operands had 0 or 5 as a unit digit (e.g., 40+65), (b) no digits were repeated within operands (e.g., 22+63), (c) the first operand was larger than the second operand in half the problems (e.g., 73+38) and vice versa in the other problems (e.g., 27+64), and (d) the sum of the unit digits was never equal to 10 (e.g., 26+64).

Two other sets of 16 problems each were tested under the one-strategy condition. These problems were the same as those tested under the two-strategy condition, but the order of operands was reversed (e.g., 31+58 was in one of the first two sets of problems tested under the two-strategy condition and 58+31 was in one of the two sets of problems tested under the one-strategy condition). One of these two sets of 16 problems was solved with the MRD strategy and the other with the MRU strategy by half the participants, and the reverse for the other participants.

2.3 Procedure

Before encountering the experimental problems, children were told that they were going to do computational estimation. The computational estimation task was explained as giving an approximate answer to an arithmetic problem (e.g., 34+57) that is as close as possible to the correct answer without actually calculating the correct answer. They were told to use only two rounding strategies, the MRD or the MRU strategy. With the MRD strategy, participants had to round the first operand down and the second operand up to their closest decades (e.g., 30+60). With the MRU strategy, participants had to round the first operand up and the second operand down to their closest decades (e.g., 40+50). A cue, which was presented above the problem, indicated which strategy participants had to use (i.e., OB for Down Up in Dutch served as a cue for the MRD strategy and BO for Up Down in Dutch served as a cue for the MRU strategy). Participants had to use each strategy on half of the problems, whereby for half of the problems each strategy was the "better" strategy (e.g., when participants had to use the MRD strategy on problems with the unit digit of the first operand smaller than 5 and that of the second operand larger than 5) and for the other half the "poorer" strategy (e.g., when participants had to use the MRD strategy on problems with the unit digit of the first operand larger than 5 and that of the second operand smaller than 5).

Before the experiment started in earnest, participants practiced on 12 problems to familiarize themselves with the two available strategies and with the procedure. Then, they solved two blocks of 16 problems each under the two-strategy condition in which the MRD or MRU strategy was randomly cued on each problem, with the constraint that participants had to switch strategy on two consecutive problems for half the items and to repeat the same strategy for the other items. Then, children solved two blocks of 16 problems each under the

one-strategy condition where respectively the MRD and MRU strategy were applied on all 16 problems. The two-strategy condition was always presented first in order to avoid carry-over effects from repeatedly applying a particular strategy in the one-strategy condition on children's strategy execution in the two-strategy condition (Siegler & Lemaire, 1997). The order in which the two strategies had to be applied in the one-strategy condition was counterbalanced across participants.

Each trial started with a 500-ms blank screen before a 400-ms fixation cross displayed at the center of the computer screen. Then, the problem was displayed and remained until participants' response. Children were asked to calculate out loud so as to be sure of which strategy they used. On each trial, the experimenter recorded children's response and strategy used. Following previous studies using this procedure (e.g., Lemaire & Brun, 2014, 2016; Hinault, Lemaire, & Phililips, 2016; Lemaire & Hinault, 2014; Uittenhove & Lemaire, 2012, 2013), timing of each response began when the problem appeared on the screen and ended when the experimenter pressed the left mouse button, the latter event occurring as soon as possible after the participant's responses.

Participants were individually tested in one session that lasted approximately 45 minutes. Between two blocks, participants were allowed a short break. The computational estimation task was run on a DELL Latitude laptop and was controlled by E-prime software (Schneider, Eschman, & Zuccolotto, 2002), which recorded latencies to the millisecond.

3. Results

Mean correct solution latencies and percentages of errors (an error estimate was coded 1 if the estimated sum differed from the expected estimate given the cued strategy, as children always used the cued strategy) were analyzed with mixed-design ANOVAs, 2 (Group: fifth, seventh graders) x 2 (Condition: one-strategy, two-strategy) x 2 (Strategy: mixed rounding-down, mixed rounding-up), with repeated measures on the two last factors (see Table 1). In all results, unless otherwise noted, differences are significant to at least $p<.05$.

Seventh graders were faster than fifth graders (4935 ms vs. 5756 ms; $F(1,70)=11.02$, $MSe=4411337.0$, $\eta^2_p=0.14$). All children were faster in the one-strategy condition (5094 ms) than in the two-strategy condition (5597 ms; $F(1,70)=59.02$, $MSe=308942.0$, $\eta^2_p=0.46$) and executed the MRD-strategy (5258 ms) more quickly than the MRU-strategy (5433 ms; $F(1,70)=11.51$, $MSe=192379.8$, $\eta^2_p=0.14$). The Group x Condition interaction ($F(1,70)=5.49$, $MSe=308942.0$, $\eta^2_p=0.07$) revealed that the effect of condition was larger in fifth graders (657 ms) than in seventh graders (350 ms). The Condition x Strategy interaction was significant ($F(1,70)=28.10$, $MSe=181469.1$, $\eta^2_p=0.29$), and the Group x Condition x Strategy was marginally significant ($F(1,70)=3.34$, $MSe=181469.1$, $\eta^2_p=0.05$, $p=.07$).

Separate analyses in each age group revealed that the Condition x Strategy interaction was significant in fifth graders ($F(1,35)=18.65$, $MSe=247674.5$, $\eta^2_p=0.35$) and in seventh graders ($F(1,35)=9.50$, $MSe=115263.6$, $\eta^2_p=0.21$). Fifth graders were 531-ms ($F(1,35)=15.51$, $MSe=327242.6$, $\eta^2_p=0.31$) faster with the MRD strategy than with the MRU strategy in the one-strategy condition and 185-ms ($F(1,35)=4.15$, $MSe=147609.2$, $\eta^2_p=0.11$) faster with the MRU than with the MRD strategy in the two-strategy condition. Seventh graders were 352-ms ($F(1,35)=11.90$, $MSe=187398.3$, $\eta^2_p=0.25$) faster with the MRD strategy than with the MRU strategy under the one-strategy condition but were equally fast with the MRD and MRU strategies in the two-strategy condition ($F<0$).

Table 1. Mean solution latencies (in ms) and percentages of errors for each strategy in each group under two-strategy and one-strategy conditions

Strategies	Fifth Graders				Seventh Graders			
	One-Strategy	Two-Strategy	*Means*	*Differences*	One-Strategy	Two-Strategy	*Means*	*Differences*
	Latencies (in ms)							
MRD Strategy	5163	6177	5670	1015	4584	5108	4846	524
MRU Strategy	5694	5992	5843	299	4936	5111	5024	175
Means	5428	6085	5756	657	4760	5110	4935	350

Differences	531	-185	173		352	3	178	
				Percentages of Errors				
MRD Strategy	1.1	1.1	1.1	0.0	2.8	0.3	0.4	-2.4
MRU Strategy	0.5	2.4	2.6	1.9	0.0	0.0	0.0	0.0
Means	0.8	1.8	1.3	1.0	0.3	0.2	0.2	-0.1
Differences	-0.6	1.4	1.5		-2.8	-0.3	-0.4	

Note. MRD: Mixed Rounding-Down strategy; MRU: Mixed Rounding-Up strategy.

As can be seen in Table 1, children were very accurate. They erred on average on less than 2% of the problems. Fifth graders made more errors than seventh graders (1.8% vs. 0.2%; $F(1,70)=14.41$, $MSe=13.1$, $\eta^2_p=0.17$), and all children tended to err more while executing the MRD strategy compared to the MRU strategy (0.8% vs. 1.3%, $F(1,70)=3.17$, $MSe=6.5$, $\eta^2_p=0.04$, $p=.08$). The Group x Strategy interaction ($F(1,70)=10.48$, $MSe=6.5$, $\eta^2_p=0.13$) revealed that this strategy difference was larger in fifth graders (1.5%) than in seventh graders (-0.4%). No other effects came out significant (Fs<1).

4. General Discussion

In this study, we investigated how strategy execution changes with children's age and the number of available strategies. Fifth and seventh graders accomplished computational estimation task in conditions where only one versus two strategies were available. To control for strategy selection, which strategy had to be executed was cued for each problem. Results showed that children were surprisingly accurate. Due to these very low error rates, the accuracy measure was most probably not sufficiently sensitive to detect reliable differences between the one- and the two-strategy condition. With respect to speed, we found that both groups of children were faster under the one-strategy condition than under the two-strategy condition, and that group-related differences were larger under the two-strategy condition. Also, differences in speed tended to vary as a function of the number of strategies, and this strategy difference was largest in younger children. These findings have implications to further our understanding of strategy execution in arithmetic and in other cognitive domains, as well as of age-related differences in children's performance during cognitive development.

Although in both two- and one-strategy conditions participants did not have to select which strategy to use on each problem, the two-strategy condition differed from the one-strategy condition in several respects. First, the two available strategies were maintained active in the two-strategy condition, as either of them could be unpredictably cued on each problem. Such active maintenance may have consumed resources unavailable for most efficient execution of the cued strategy. In contrast, all processing resources were available for strategy execution in the one-strategy condition. Also, when participants encoded the to-be-executed strategy and the problem, they had to (at least partially) inhibit the irrelevant strategy and activate procedures of the cued strategy before executing them. Such inhibitory processes are not involved in the one-strategy condition. Finally, switching strategies across successive problems may have used resources that could not be used for strategy execution. Note though that we found no strategy switch costs, as seen in comparable latencies when children used the same strategy versus different strategies on two successive problems (5637 ms vs. 5528 ms, F<1). This lack of strategy switch costs contrasts with previous findings (e.g., Lemaire & Leacheur, 2010; Lemaire & Brun, 2014). Although it is possible that switching between strategies incurs no switching costs, in contrast to what has been found previously, it is also possible that switching between strategies incurred costs in this study but durations of switching processes were absorbed by durations of other processes involved during strategy execution (note that average solution times were over 4900 ms). Nevertheless, a number of differences between two- and one-strategy conditions may have contributed to strategies being executed more slowly under the two-strategy condition.

Interestingly, we found that relative strategy performance differed across two- and one-strategy conditions. Children were 441-ms faster with the MRD than with the MRU strategy under the one-strategy condition and equally fast with both strategies under the two-strategy condition. MRD is easier than MRU, possibly because once the first operand is rounded down, this rounded operand is stored in working memory before adding it to the second operand. Executing MRU involves first rounding the second operand up and storing this first rounded operand in working memory, then rounding the second operand down while maintaining the first operand in

working memory, and finally adding both rounded operands. Executing MRD involves first rounding first operand down and storing this rounded operand in working memory, then rounding second operand up, and finally adding both rounded operands. In other words, adding one rounded operand to the other operand takes more time if the rounding operation is carried out after encoding the first operand and temporarily storing it in working memory than if the rounding operation is first carried out before adding the second operand to this rounded operand.

Differences in strategy performance under the one-strategy condition and the lack of such differences in the two-strategy condition suggest that one-strategy condition yields better assessment of strategy execution (i.e., uncontaminated by other factors). Indeed, in the two-strategy condition, although relative strategy performance is not contaminated by strategy distribution and strategy selection (as both strategies are used by all participants on an equal number of problems), differences in strategy performance may have been absorbed by extra-time required to manage two strategies as is suggested by longer latencies in the two-strategy condition than in the one-strategy condition. Thus, as already argued by Siegler and Lemaire (1997) and many others since, when we want to assess relative difficulty of strategy execution and compare strategy performance, it is best to test participants under a condition in which all participants have to execute the same strategy on all problems, and to do this for all available strategies.

The final set of findings of interest in this experiment concerns age-related differences in children's performance. It is possible that these differences could partially be explained by age-related differences in experience with and skills underlying these strategies. However, as these age-related differences were modulated by the type of strategies, other more domain-general factors, such as cognitive resources, might contribute to these differences as well. Age-related differences in performance were larger under the two-strategy condition than under the one-strategy condition, as increased latencies from the one- to the two-strategy condition were larger in younger children. Again managing two strategies incurred more cognitive resources and as young children have fewer resources available, they increased their latencies in the two-strategy condition relative to the one-strategy condition to a larger extent than older children.

Both fifth and seventh graders were faster with the MRD strategy than with the MRU strategy. However, this strategy difference was found only in the one-strategy condition where it was larger in third than in fifth graders. This reflects typical larger age differences on the more difficult strategy that has been found in a number of cognitive domains (see Siegler, 1996). As the harder strategy requires more processing resources to execute and younger children have fewer resources available, they need more time to execute the harder strategy.

Even though previous studies have already demonstrated that the number of strategies that individuals are using is related to their executive resources (Ardiale et al., 2012; Hodzik & Lemaire, 2011), future studies should examine the degree to which the effects observed in the present study could be attributed to children's measures of executive functions, like inhibition and shifting capacities, as well as working-memory. This approach will provide a direct test of the extent to which the number of strategies in someone's repertoire consumes executive resources and working-memory capacity.

The present findings have some potential educational implications as it is shown that the number of available strategies negatively affects children's arithmetic performance and that this effect becomes smaller with age. Although accuracy is a more important performance measure in education contexts, and we observed this effect only on children's latencies and not on their error rates, it cannot be ruled out that similar effects would occur on children's accuracy as well. Indeed, as mentioned earlier, children were surprisingly very accurate. Including more difficult problems in future work might exclude such ceiling effects and reveal similar effects on accuracy. In that case, it might be recommendable to instruct and let practice young children one strategy at a time, as to spare their cognitive resources and working-memory capacity. At a later age, when children's executive functions and working-memory capacity have improved, one might consider teaching multiple strategies at the same time (see also Jitendra et al., 2007).

The present effects of the number of strategies have important theoretical implications to further our understanding of how children execute strategies and age-related differences therein. Formal models of strategies (e.g., Lovett & Anderson, 1996, ACT-R model; Lovett & Schunn, 1999, RCCL model; Payne et al., 1993, adaptive decision maker model; Rieskamp & Otto, 2006, SSL model; or Siegler & Arraya, 2005, SCADS* model) share core assumptions regarding how participants execute strategies on each problem. For example, all models proposed that strategy performance depends on the number and difficulty of the different processing steps within a strategy. These models also assume that how each procedure within a strategy is executed depends on problem, situation,

and participants characteristics. Finally, these models assume that age-related changes in children's speed of strategy execution involves an increase in the speed of triggering and executing the different component processes within a strategy. In other words, strategy execution on a given problem and age-related changes in strategy execution are, according to these models, independent of the number of strategies that are used across all problems in a given task. The effects of the number of strategies found here suggest that strategy execution and age-related changes in how children execute strategies are also influenced by processes that enable managing several strategies. Such processes involve working-memory and executive control processes (e.g., inhibition, switching). Assumptions of current models of strategies could be augmented to include assumptions regarding the role of these processes during strategy execution. Moreover, additional assumptions in current models of strategy selection and execution could computationally specify how relative strategy performance becomes more and more independent of the number of strategies being available for solving the different problems in a given task and, more generally, how developmental mechanisms enable children to more and more efficiently manage several strategies.

Acknowledgments

This work was supported by a grant from La Fondation de France to PL, under Grant # BLAN-1912-01 to PL and the grant C16/16/001 of the Research Fund KU Leuven to KL. We would like to thank Katrien Franssen and Helena Van Hoof for their help during data collection.

References

Ardiale, E., Hodzik, S., & Lemaire, P. (2012). Aging and strategy switch costs: A study in arithmetic problem solving. *L'Année Psychologique*, *112*, 345-360. https://doi.org/10.4074/S0003503312003028

Campbell, J. I. D. (2005). *Handbook of mathematical cognition*. New York: Psychology Press.

Cohen Kadosh, R., & Dowker, A. (2015). *The Oxford handbook of numerical cognition*. New York: Oxford University Press.

Dowker, A. (1997). Young children's addition estimates. *Mathematical Cognition*, *3*(2), 141-154. https://doi.org/10.1080/135467997387452

Dowker, A., Flood, A., Griffiths, H., Harris, L., & Hook, L. (1996). Estimation strategies of four groups. *Mathematical Cognition*, *2*(2), 113-135. https://doi.org/10.1080/135467996387499

Geary, D. C. (1994). *Children's mathematical development*. Washington, DC: American Psychological Association.

Hinault, T., Lemaire, P., & Phillips, N. (2016). Aging and sequential modulations of poorer-strategy effects: An EEG study in arithmetic problem solving. *Brain Research*, *1630*, 144-158. https://doi.org/10.1016/j.brainres.2015.10.057

Hodzik, S., & Lemaire, P. (2011). Inhibition and shifting capacities mediate adults' age-related differences in strategy selection and repertoire. *Acta Psychologica*, *137*, 335-344. https://doi.org/10.1016/j.actpsy.2011.04.002

Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A comparison of single and multiple strategy instruction on third-grade students' mathematical problem solving. *Journal of Educational Psychology*, *99*(1), 115-127. https://doi.org/10.1037/0022-0663.99.1.115

LeFevre, J., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. *Cognition and Instruction*, *11*(2), 95-132. https://doi.org/10.1207/s1532690xci1102_1

Lemaire, P., & Brun, F. (2014). Effects of strategy sequences and response-stimulus intervals on children's strategy selection and strategy execution: A study in computational estimation. *Psychological Research*, *78*(4), 506-519. https://doi.org/10.1007/s00426-013-0501-0

Lemaire, P., & Brun, F. (2016). Age-related differences in children's strategy repetition: A study in arithmetic. *Journal of Experimental Child Psychology*, *150*, 227-240. https://doi.org/10.1016/j.jecp.2016.05.014

Lemaire, P., & Hinault, T. (2014). Age-related differences in sequential modulations of poorer-strategy effects: A study in arithmetic problem solving. *Experimental Psychology*, *61*(4), 253-262. https://doi.org/10.1027/1618-3169/a000244

Lemaire, P., & Lecacheur, M. (2002). Children's strategies in computational estimation. *Journal of Experimental Child Psychology, 82*, 281-304. https://doi.org/10.1016/S0022-0965(02)00107-8

Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition, 38*(3), 322-332. https://doi.org/10.3758/MC.38.3.322

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development, 26*(3), 282-294. https://doi.org/10.1016/j.cogdev.2011.01.002

Lemaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic? An example of parity and five effects on product verification. *Memory & Cognition, 22*, 364-382. https://doi.org/10.3758/BF03211420

Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology, 31*(2), 168-217. https://doi.org/10.1006/cogp.1996.0016

Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability, and base-rate neglect. *Journal of Experimental Psychology: General, 128*(2), 107-130. https://doi.org/10.1037/0096-3445.128.2.107

Marewski, J. N., & Link, D. (2014). Strategy selection: An introduction to the modeling challenge. *WIREs Cognitive Sciences, 5*, 39-59. https://doi.org/10.1002/wcs.1265

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139173933

Reitan, R., & Wolfson, D. (1992). *Neuropsychological evaluation of older children.* Tucson, AZ: Neuropsychology Press.

Reys, R., Bestgen, B., Rybolt, J., & Wyatt, J. (1982). Processes used by good computational estimators. *Journal for Research in Mathematics Education, 12*(3), 183-201. https://doi.org/10.2307/748555

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*(2), 207-236. https://doi.org/10.1037/0096-3445.135.2.207

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Siegler, R. S. (1996). *Emerging minds*. New York: Oxford University Press.

Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10*, 104-209. https://doi.org/10.1111/j.1467-7687.2007.00571.x

Siegler, R. S., & Arraya, R. (2005). A computational model of conscious and unconscious strategy discovery. In R. V. Kail (Ed.), *Advances in child development and behaviour* (Vol. 33, pp. 1-42). Oxford, UK: Elsevier. https://doi.org/10.1016/s0065-2407(05)80003-5

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General, 126*, 71-92. https://doi.org/10.1037/0096-3445.126.1.71

Sowder, J. T., & Markovits, Z. (1990). Relative and absolute error in computational estimation. In G. Booker, P. Cobb, & T. N. de Mendicuti (Eds.), *Proceedings of the Fourteenth Psychology of Mathematics Education Conference* (pp. 321-328). Mexico.

Uittenhove, K., & Lemaire, P. (2012). Sequential difficulty effects during strategy execution: A study in arithmetic. *Experimental Psychology, 59*(5), 295-301. https://doi.org/10.1027/1618-3169/a000157

Uittenhove, K., & Lemaire, P. (2013). Strategy sequential difficulty effects vary with working-memory and response-stimulus intervals: A study in arithmetic. *Acta Psychologica, 143*, 113-118. https://doi.org/10.1016/j.actpsy.2013.02.007

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Parents' Social Validity Appraisals of Early Childhood Intervention Practice Guides

Carl J. Dunst[1]

[1] Orelena Hawks Puckett Institute, Asheville and Morganton, NC, USA

Correspondence: Carl J. Dunst, Orelena Hawks Puckett Institute, 128 S. Sterling Street, Morganton, NC, 28655, USA. Tel: 1-828-432-0065. E-mail: cdunst@puckett.org

Received: June 5, 2017 Accepted: June 15, 2017 Online Published: July 4, 2017

doi:10.5539/jedp.v7n2p51 URL: http://doi.org/10.5539/jedp.v7n2p51

Abstract

Findings from three field-tests of parents' ratings of early childhood intervention practice guides are reported. Results from the first field-test were used to inform changes to the practice guides in the second field-test, and results from the second field-test were used to inform changes to the practice guides in the third field-test. Parents' judgments of the practice guide designs and their social validity appraisals of the practice guide intervention activities and child outcomes were correlated with parent-informed improvements in the intervention materials. The results add to the knowledge base in terms of how parent-informed improvements to the practice guides are related to product design judgments and the social validity appraisals of the importance and acceptability of early childhood intervention materials.

Keywords: early childhood intervention, parent practice guides, product design judgments, social validity appraisals

1. Introduction

Early childhood intervention includes the learning experiences and opportunities used with infants, toddlers, and preschoolers to promote and enhance the children's development (Dunst & Espe-Sherwindt, 2017; Groark, Eidelman, Maude, & Kaczmarek, 2011; Guralnick, 2016; McWilliam, 2015). Early childhood intervention practitioners often support and strengthen parents' and other primary caregivers' abilities to use these experiences and opportunities to promote the learning and development of their young children (e.g., Acar & Akamoğlu, 2014; Bernheimer & Keogh, 1995; Friedman, Woods, & Salisbury, 2012). There is, however, considerable variability in the extent to which parents use the experiences and opportunities as early childhood intervention practices with their children (Halgunseth, 2009; Korfmacher et al., 2008; Roggman et al., 2016).

Among the many reasons parents use or do not use early childhood intervention practices with their children are their beliefs about the importance and acceptability of the practices and the intended child outcomes of the practices (e.g., Dunst, Trivette, Prior, Hamby, & Embler, 2013b; Reimers & Wacker, 1988). These types of beliefs or subjective judgments have been described as social validity appraisals (Foster & Mash, 1999). As noted by Strain, Barton, and Dunlap (2012), an intervention practice is not likely to be used by a parent (or a practitioner) if it is not considered worth the time and effort to use. This was demonstrated in a study by Dunst, Raab, and Hamby (2016) where parents' social validity appraisals of interest-based child language learning practices were related to the fidelity of use of the practices, where fidelity of use of the practices in turn was related to improvements in the children's language development.

Cognitive appraisals of the design characteristics of materials or products also influences people's beliefs about the usability of different products and materials (Bloch, 1995). This has been described as the aesthetics-usability effect (Lidwell, Holden, & Butler, 2003). Research on the design properties (attractiveness, appearance, organization, aesthetics, etc.) of a product indicates that design matters a great deal when a person does or does not judge products, materials, etc., as having personal benefit or usability (e.g., Hamborg, Hülsmann, & Kaspar, 2014; Seva, Gosiaco, Santos, & Pangilinan, 2010; Sonderegger & Sauer, 2010; Spague, Pennefather, Marquez, Yeaton, & Marquez, 2011).

The study described in this paper is part of a line of research and practice investigating parents' and practitioners' judgments and appraisals of early childhood intervention practice guides where end-user suggestions and feedback were used to improve the design, organization, and content of the intervention materials (e.g., Dunst, 2017; Dunst & Hamby, 2017; Dunst, Trivette, Prior, Hamby, & Embler, 2013a; Trivette, Dunst, Masiello, Gorman, & Hamby, 2009). This brief report includes findings from three field-tests, where results from each field-test were used to inform changes and improvements in next generation practice guides.

Parents in all three field-tests selected from among a list of practice guides to review and make social validity appraisals of the importance and acceptability of both the practice guide intervention activities and child outcomes. The parents also made judgments of the attractiveness and appeal of the practice guide design. The progressive changes and improvements in the practice guides made in response to parent feedback were expected to be related to increases in parents' judgments of the importance, acceptability, and design features of the intervention materials. Exploratory structural equation modeling was used to trace the pathways of influence from improvements to the practice guides to the judgments of the attractiveness and appeal of the interventions materials to parents' social validity appraisals of the practice guide intervention activities and then to the social validity appraisals of the practice guide child outcomes.

2. Method

2.1 Participants

The participants were 173 parents and other primary caregivers of infants, toddlers, and preschoolers involved in early childhood intervention programs throughout the United States. The parents were recruited through parent organizations and early childhood programs and by announcements on parent organization websites and listserves.

The participants' children were receiving early childhood intervention because of identified disabilities, developmental delays, or because they were at-risk for poor developmental outcomes for medical (e.g., low birth weight) or socio-environmental (lower family socioeconomic status) factors. Forty-two percent of the children were involved in birth to age three early intervention or Early Head Start programs, and 58% of the children were involved in age three to five year old early childhood-special education or Head Start programs.

2.2 Practice Guides

The practice guides are formatted in similar ways. Each practice guide includes a description of an intervention practice and the intended benefits or outcomes of the practice, a description or list of activities (methods, strategies, etc.) for a parent to use a practice with his or her child, a vignette of parents using a practice with their children, and a list of three outcome indicators for determining if the practice guide activities had intended child benefits. Lessons learned from the first field-test informed changes and improvements in the practice guides in the second field-test, and lessons learned from both the first and second field-tests informed changes and improvements in the practice guides in the third field-test. The latter involved captioned videos of parents using the practices with their children and an external link to other resources for using the practice to promote child learning and development. Figure 1 shows one of the practice guides evaluated by parents in the third field-test.

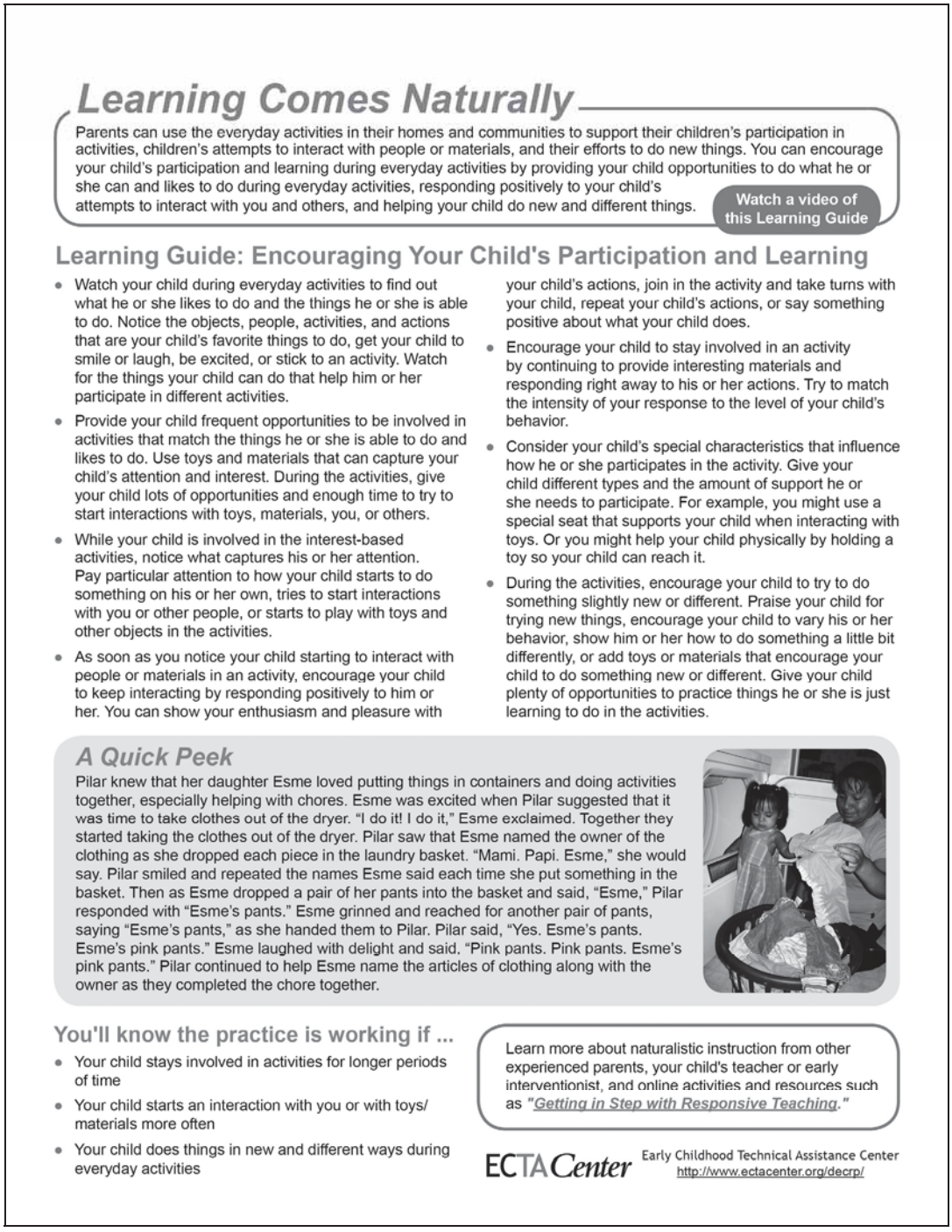
2.3 Procedure

Invitations to participate in the field-tests were posted on parent and early child intervention organization websites, on parent organization listserves, and parent organization web-based newsletters. Requests were also sent to early childhood program directors who were asked to distribute the invitations to interested parents.

The invitations included a description of the purpose of the field-tests and instructions for choosing and reviewing a practice guide. The respondents were asked to read the practice guide with a specific focus on the practice guide activities and intended child outcomes. The invitation also included a web-based link to the field-test survey for respondents to evaluate the practice guides.

The survey included eight parent appraisal items, space for making comments or suggestions for improving the practice guides, a question about the age of the parent's child, and a question about their child's early childhood intervention program. The survey included items measuring parents' appraisals of the practice guide design, practice guide intervention activities, and practice guide child outcomes. The parent appraisal items included four items measuring the social validity of the practice guide intervention activities (e.g., "*The practice would easily fit into my everyday schedule*"; "*The practice would be worth my time and effort to use*"), three social validity items measuring the intended child outcomes of the practice (e.g., "*The practice would help my child*

learn"; "*The practice guide activities would be interesting to my child*"), and one program design item ("*The practice guide format is both attractive and appealing*"). The eight items were each rated on a 5-point scale ranging from *Do Not Agree at All* to *Agree a Great Deal* with the survey statements.



Learning Comes Naturally

Parents can use the everyday activities in their homes and communities to support their children's participation in activities, children's attempts to interact with people or materials, and their efforts to do new things. You can encourage your child's participation and learning during everyday activities by providing your child opportunities to do what he or she can and likes to do during everyday activities, responding positively to your child's attempts to interact with you and others, and helping your child do new and different things.

Watch a video of this Learning Guide

Learning Guide: Encouraging Your Child's Participation and Learning

- Watch your child during everyday activities to find out what he or she likes to do and the things he or she is able to do. Notice the objects, people, activities, and actions that are your child's favorite things to do, get your child to smile or laugh, be excited, or stick to an activity. Watch for the things your child can do that help him or her participate in different activities.
- Provide your child frequent opportunities to be involved in activities that match the things he or she is able to do and likes to do. Use toys and materials that can capture your child's attention and interest. During the activities, give your child lots of opportunities and enough time to try to start interactions with toys, materials, you, or others.
- While your child is involved in the interest-based activities, notice what captures his or her attention. Pay particular attention to how your child starts to do something on his or her own, tries to start interactions with you or other people, or starts to play with toys and other objects in the activities.
- As soon as you notice your child starting to interact with people or materials in an activity, encourage your child to keep interacting by responding positively to him or her. You can show your enthusiasm and pleasure with your child's actions, join in the activity and take turns with your child, repeat your child's actions, or say something positive about what your child does.
- Encourage your child to stay involved in an activity by continuing to provide interesting materials and responding right away to his or her actions. Try to match the intensity of your response to the level of your child's behavior.
- Consider your child's special characteristics that influence how he or she participates in the activity. Give your child different types and the amount of support he or she needs to participate. For example, you might use a special seat that supports your child when interacting with toys. Or you might help your child physically by holding a toy so your child can reach it.
- During the activities, encourage your child to try to do something slightly new or different. Praise your child for trying new things, encourage your child to vary his or her behavior, show him or her how to do something a little bit differently, or add toys or materials that encourage your child to do something new or different. Give your child plenty of opportunities to practice things he or she is just learning to do in the activities.

A Quick Peek

Pilar knew that her daughter Esme loved putting things in containers and doing activities together, especially helping with chores. Esme was excited when Pilar suggested that it was time to take clothes out of the dryer. "I do it! I do it," Esme exclaimed. Together they started taking the clothes out of the dryer. Pilar saw that Esme named the owner of the clothing as she dropped each piece in the laundry basket. "Mami. Papi. Esme," she would say. Pilar smiled and repeated the names Esme said each time she put something in the basket. Then as Esme dropped a pair of her pants into the basket and said, "Esme," Pilar responded with "Esme's pants." Esme grinned and reached for another pair of pants, saying "Esme's pants," as she handed them to Pilar. Pilar said, "Yes. Esme's pants. Esme's pink pants." Esme laughed with delight and said. "Pink pants. Pink pants. Esme's pink pants." Pilar continued to help Esme name the articles of clothing along with the owner as they completed the chore together.

You'll know the practice is working if ...

- Your child stays involved in activities for longer periods of time
- Your child starts an interaction with you or with toys/ materials more often
- Your child does things in new and different ways during everyday activities

Learn more about naturalistic instruction from other experienced parents, your child's teacher or early interventionist, and online activities and resources such as "*Getting in Step with Responsive Teaching*."

ECTA Center Early Childhood Technical Assistance Center
http://www.ectacenter.org/decrp/

Figure 1. Example of a practice guide that was the focus of parents' evaluative judgments and appraisals

2.4 Methods of Analysis

Three between field-test ANOVAs with *a priori* tests for linear trends were used to determine if there were progressive increases in both the parents' practice guide design ratings and social validity appraisals of the practice guide intervention activities and child outcomes. The dependent variables were the average respondent ratings for the three sets of items so that the scores for all three measures ranged between 1 and 5. Cohen's *d* effect sizes for the linear trends were the primary metrics for substantive interpretation of the results since effect sizes rather than *p*-values provide the best estimate of the magnitude of improvements to the practice guide design and content (Coe, 2002).

The fit of the hypothesized model to the pattern of relationships among the field-test variables in the exploratory structural equation model were evaluated by the Root Mean Square Error of Approximation (RMSEA), Root Mean Square Residual (RMSR), Comparative Fit Index (CFI), and Incremental Fit Index (IFI). The closer RMSEA and RMSR are to zero, and the closer CFI and IFI are to one, the better the fit of the model to the data. The standardized structural (path) coefficients were used to evaluate the direct and indirect effects of the variables in the model. These can range from -1.00 to 1.00 when the size of effect is a measure of the strength of relationships among the variables in the model.

3. Results

3.1 Between Field-Test Comparisons

The mean ratings for the practice guide design judgments and social validity appraisals are shown in Figure 2. There were between field-test differences in the mean scores for the practice guide design, $F(2, 171) = 8.16$, $p = .0004$. There was also a linear increase in the parents' judgment of the practice guide design, $F(1, 171)$, $p = .0000$, Cohen's $d = 1.71$. Follow-up effect size calculations showed that the linear increases were primarily for the Field-Test 1 vs. Field-Test 2 and the Field-Test 1 vs. Field-Test 3 differences. The Cohen's *d* effect sizes for these two field-test comparisons were $d = .64$ and $d = .80$, respectively, for the parents' judgments of the practice guide design, whereas the effect size for Field-Test 2 vs. Field-Test 3 was $d = .21$.

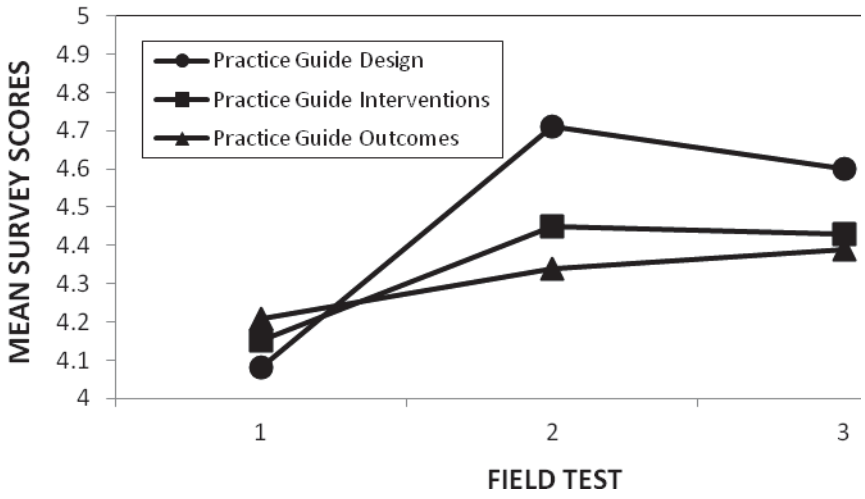


Figure 2. Parents' mean practice guide design judgments and social validity appraisals of the practice guide intervention activities and child outcomes

There were also between field-test differences in the mean scores for the parents' social validity appraisals of the practice guide intervention activities, $F(2, 171) = 3.38$, $p = .0363$, and a linear increase in the social validity scores, $F(1, 170) = 5.94$, $p = .0158$, $d = 1.02$. Follow-up effect size calculations showed that the effect sizes for the parents' social validity appraisals of the practice guide intervention activities was $d = .46$ for Field-Test 1 vs. Field-Test 2 and $d = .49$ for Field-Test 1 vs. Field-Test 3. In contrast, the effect sizes for Field-Tests 2 vs. 3 was $d = .04$.

There were no between field-test differences for the parents' social validity appraisals of the practice guide child outcomes, $F(2, 171) = 0.68$, $p = .5090$. There was, however, a small effect size for the linear change in the mean scores, $F(1, 170) = 1.31$, $p = .2540$, $d = .48$.

3.2 Correlational Analyses

Table 1 shows the correlations between the field-test study measures. The patterns of relationships were as expected. Improvements to the practice guides were correlated with both the practice guide design judgments and social validity appraisals of the practice guide intervention activities but not the child outcomes. Practice guide design judgments were correlated with both social validity appraisals, and the social validity appraisals of the practice guide intervention activities were correlated with social validity appraisals of the child outcomes.

Table 1. Correlations among the field-test variables

Study Variables	PG Design Judgments	Social Validity Appraisals	
		PG Activities	PG Outcomes
Practice Guide (PG) Improvements	.28**	.18*	.08
Practice Guide Design	-	.59***	.49***
PG Intervention Activities		-	.68***
PG Child Outcomes			-

*$p < .001$. **$p < .0002$. ***$p < .0001$.

3.3 Structural Equation Modeling Results

Figure 3 shows the results from the structural equation modeling analysis. RMSEA was .03, RMSR was .02, CFI was .99, and IFI was .99. These results indicate a good fit of the model to the relationships among the variables in the model (Table 1).

Progressive changes to the practice guides were directly related to parents' judgment of the practice guide designs, and indirectly related to parents' social validity appraisals of the practice guide intervention activities, mediated by practice guide design judgments, $\beta = .27 \times .57 = .15$, $p = .000$. The progressive changes to the practice guides were also indirectly related to the social validity appraisals of the practice guide child outcomes mediated by both practice guide design judgments and practice guide intervention activity social validity appraisals, $\beta = .14$, $p = .0035$.

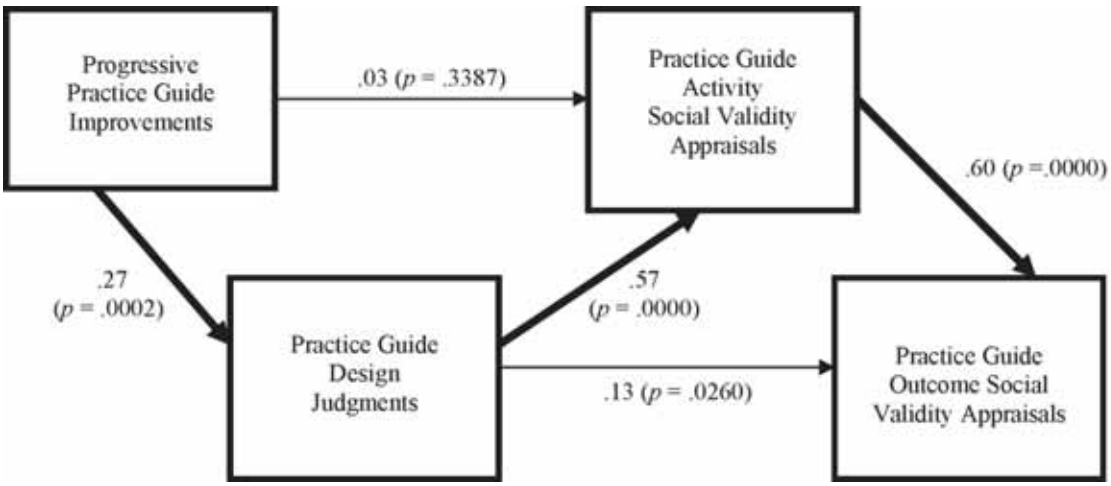


Figure 3. Pathways of relationships among the four variables in the structural equation modeling analysis

Note. The main pathways in the model are in bold.

Parents' practice guide design judgments were directly related to parents' social validity appraisals of both the practice guide intervention activities and child outcomes, although the size of effect for the relationship with the intervention activities was four times larger than those for the child outcomes. The parents' practice guide design judgments were also indirectly related to the parents' social validity appraisals of the practice guide child outcomes mediated by the social validity appraisals of the practice guide intervention activities, $\beta = .57 \times .60 = .34$, $p = .0000$. The parents' social validity appraisals of the practice guide intervention activities were directly related to their social validity appraisals of the practice guide child outcomes.

4. Discussion

Findings from the between field-test comparisons showed that parents' product design judgments and social validity appraisals of the practice guide intervention activities and child outcomes increased as a function of improvements to the intervention materials. Results from the structural equation modeling analyses showed that there was an adequate fit of the hypothesized model to the pattern of relationships among the field-test variables. The two sets of findings, taken together, indicate that parent-informed improvements in early childhood intervention materials in general, and the parent practice guides specifically, can enhance the usability, acceptability, and importance of intervention products, materials, and practices.

As noted earlier, the field-tests described in this paper are part of a line of research and practice investigating (a) improvements in both parent and practitioner early childhood intervention practices (Dunst, 2017; Dunst, Pace, & Hamby, 2007; Dunst, Trivette, Gorman, & Hamby, 2010), (b) the relationship between end-user appraisals and judgments of the practices and fidelity of use of the practices (e.g., Dunst & Hamby, 2015; Dunst, Trivette, & Raab, 2014), and (c) the effects of fidelity of use of early childhood intervention practices on child outcomes (Dunst et al., 2016). This study adds to this knowledge base by demonstrating that end-user appraisals and judgments of intervention practices are related in discernible ways. The structural equation modeling results, together with those found in other studies (e.g., Dunst et al., 2016; Dunst et al., 2014; Trivette, Raab, & Dunst, 2014), also add to the knowledge base in terms of an understanding of the manner in which different types of personal beliefs influence judgments of materials and products (e.g., Dunst & Hamby, 2017; Seva et al., 2010).

Enhancing the usability of intervention materials (Santos, Fowler, Corso, & Bruns, 2000) and the acceptance and importance of intervention practices (Strain et al., 2012) have been "called for" to improve parents' and practitioners' adoption and use of early childhood intervention practices with infants, toddlers, and preschoolers. Santos et al. (2000), for example, noted that "we need to find ways to enhance the usability of [intervention] materials" (p. 20) so as to be acceptable to a wide range of end-users. Similarly, Strain et al. (2012) noted that improving judgments of the acceptability and importance of intervention practices can facilitate increased use of the practices. Studies like the one described in this paper contribute to these two goals by involving end-users in evaluating early childhood intervention practices and materials and making end-user-informed improvements in the practices and materials. The interested reader is referred to Rice and Valdivia (1991), Santos et al. (2000), and Springston and Champion (2004) for suggestions and guidelines for designing user friendly materials.

Acknowledgments

The field-test evaluations in this paper were supported, in part, by funding from the U.S. Department of Education, Office of Special Education Programs (Grant # H326P120002) awarded to the University of North Carolina-Chapel Hill. The opinions expressed, however, are those of the authors and endorsement by the funder or university should not be inferred or implied.

References

Acar, S., & Akamoğlu, Y. (2014). Practices for parent participation in early intervention/early childhood special education. *International Journal of Early Childhood Special Education*, *6*(1), 80-101. https://doi.org/10.20489/intjecse.93010

Bernheimer, L. P., & Keogh, B. K. (1995). Weaving interventions into the fabric of everyday life: An approach to family assessment. *Topics in Early Childhood Special Education*, *15*, 415-433. https://doi.org/10.1177/027112149501500402

Bloch, P. H. (1995). Seeking the ideal form: Product design and consumer response. *Journal of Marketing, 5*(3), 16-29. https://doi.org/10.2307/1252116

Coe, R. (2002, September). *It's the effect size, stupid: What effect size is and why it is important*. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England. Retrieved from http://www.leeds.ac.uk/educol/documents/00002182.htm

Dunst, C. J. (2017). Early childhood practitioner judgments of the social validity of performance checklists and parent practice guides. *Journal of Education and Training Studies*, *5*(3), 176-187. https://doi.org/10.11114/jets.v5i3.2162

Dunst, C. J., & Espe-Sherwindt, M. (2017). Contemporary early intervention models, research, and practice for infants and toddlers with disabilities and delays. In J. M. Kauffman, D. P. Hallahan, & C. P. Pullen (Eds.), *Handbook of special education* (2nd ed., pp. 831-849). New York: Routledge.

Dunst, C. J., & Hamby, D. W. (2015). Research synthesis of studies to promote parent and practitioner use of assistive technology and adaptations with young children with disabilities. In D. L. Edyburn (Ed.), *Advances in special education technology: Efficacy of assistive technology interventions* (Vol. 1, pp. 51-78). United Kingdom: Emerald Publishing.

Dunst, C. J., & Hamby, D. W. (2017). Predictors of the social validity judgments of early childhood intervention performance checklists and practice guides. *International Journal of Psychology and Educational Studies*, *4*(1), 13-20. https://doi.org/10.17220/ijpes.2017.01.002

Dunst, C. J., Pace, J., & Hamby, D. W. (2007). *Evaluation of the Games for Growing tool kit for promoting early contingency learning*. Asheville, NC: Winterberry Press.

Dunst, C. J., Raab, M., & Hamby, D. W. (2016). Interest-based everyday child language learning. *Revista de Logopedia, Foniatria y Audiologia*, *36*, 153-161. https://doi.org/10.1016/j.rlfa.2016.07.003

Dunst, C. J., Trivette, C. M., & Raab, M. (2014). Everyday child language learning early intervention practices. *Infants and Young Children*, *27*(3), 207-219. https://doi.org/10.1097/IYC.0000000000000015

Dunst, C. J., Trivette, C. M., Gorman, E., & Hamby, D. W. (2010). Further evidence for the social validity of the Center for Early Literacy Learning practice guides. *CELLpapers*, *5*(1), 1-3. Retrieved from http://www.earlyliteracylearning.org/cellpapers/cellpapers_v5n1.pdf

Dunst, C. J., Trivette, C. M., Prior, J., Hamby, D. W., & Embler, D. (2013a). Parents' appraisals of the animacy and likeability of socially interactive robots for intervening with young children with disabilities. *Social Robots Research Reports*, *2*. Retrieved from http://www.socialrobots.org/SocRobotRpt_2.pdf

Dunst, C. J., Trivette, C. M., Prior, J., Hamby, D. W., & Embler, D. (2013b). Parents' judgments of the acceptability and importance of socially interactive robots for intervening with young children with disabilities. *Social Robots Research Reports*, *1*. Retrieved from http://www.socialrobots.org/reports.php

Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research issues and procedures. *Journal of Consulting and Clinical Psychology*, *67*, 308-319. https://doi.org/10.1037/0022-006X.67.3.308

Friedman, M., Woods, J., & Salisbury, C. (2012). Caregiver coaching strategies for early intervention providers: Moving toward operational definitions. *Infants and Young Children*, *25*(1), 62-82. https://doi.org/10.1097/IYC.0b013e31823d8f12

Groark, C. J., Eidelman, S. M., Maude, S., & Kaczmarek, L. (2011). *Early childhood intervention: Shaping the future for children with special needs and their families*. Santa Barbara, CA: Praeger.

Guralnick, M. J. (2016). Early intervention for children with intellectual disabilities: An update. *Journal of Applied Research in Intellectual Disabilities*, *30*(2), 211-229. https://doi.org/10.1111/jar.12233

Halgunseth, L. C. (2009). Family engagement, diverse families and early childhood education programs: An integrated review of the literature. *Young Children*, *64*(5), 56-58. Retrieved from http://search.proquest.com.libproxy.lib.unc.edu/docview/197647848/fulltextPDF/6AE90A99362645CEPQ/11?accountid=14244

Hamborg, K.-C., Hülsmann, J., & Kaspar, K. (2014). The interplay between usability and aesthetics: More evidence for the "what is usable is beautiful" notion. *Advances in Human-Computer Interaction* (Article ID 946239), *2014*. https://doi.org/10.1155/2014/946239

Korfmacher, J., Green, B., Staerkel, F., Peterson, C., Cook, G. A., Roggman, L. A., … Schiffman, R. (2008). Parent involvement in early childhood home visiting. *Child and Youth Care Forum*, *37*(4), 171-196. https://doi.org/10.1007/s10566-008-9057-3

Lidwell, W., Holden, K., & Butler, J. (2003). *Universal principles of design*. Beverly, MA: Rockport Publishers.

McWilliam, R. A. (2015). Future of early intervention with infants and toddlers for whom typical experiences are not effective. *Remedial and Special Education*, *36*(1), 33-38. https://doi.org/10.1177/0741932514554105

Reimers, T. M., & Wacker, D. P. (1988). Parents' ratings of the acceptability of behavioral treatment recommendations made in an outpatient clinic: A preliminary analysis of the influence of treatment effectiveness. *Behavioral Disorders*, *14*(1), 7-15.

Rice, M., & Valdivia, L. (1991). A simple guide for design, use, and evaluation of educational materials. *Health Education Quarterly*, *18*, 79-85. https://doi.org/10.1177/109019819101800108

Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V., Boyce, L. K., Christiansen, K., & Peterson, C. A. (2016). Home visit quality variations in two Early Head Start programs in relation to parenting and child vocabulary outcomes. *Infant Mental Health Journal*, *37*(3), 193-207. https://doi.org/10.1002/imhj.21565

Santos, R. M., Fowler, S. A., Corso, R. M., & Bruns, D. A. (2000). Acceptance, acknowledgment, and adaptability: Selecting culturally and linguistically appropriate early childhood materials. *Teaching Exceptional Children*, *32*(3), 14-22. https://doi.org/10.1177/004005990003200303

Seva, R. R., Gosiaco, K. G. T., Santos, C. E. D., & Pangilinan, D. M. L. (2010). Product design enhancement using apparent usability and affective quality. *Applied Ergonomics*, *30*, 1-7.

Sonderegger, A., & Sauer, J. (2010). The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied Ergonomics*, *41*(3), 403-410. https://doi.org/10.1016/j.apergo.2009.09.002

Spague, J., Pennefather, J., Marquez, J., Yeaton, P., & Marquez, B. (2011). *Online universal screening and behavioral progress monitoring: Assessing social validity, usability and intent to use by K-3 teachers*. Retrieved from https://www.sree.org/conferences/2011/program/downloads/posters/162.pdf

Springston, J. K., & Champion, V. L. (2004). Public relations and cultural aesthetics: Designing health brochures. *Public Relations Review*, *30*, 483-491. https://doi.org/10.1016/j.pubrev.2004.08.005

Strain, P. S., Barton, E. E., & Dunlap, G. (2012). Lessons learned about the utility of social validity. *Education and Treatment of Children*, *35*, 183-200. https://doi.org/10.1353/etc.2012.0007

Trivette, C. M., Dunst, C. J., Masiello, T., Gorman, E., & Hamby, D. W. (2009). Social validity of the Center for Early Literacy Learning parent practice guides. *CELLpapers*, *4*(1), 1-4. Retrieved from http://www.earlyliteracylearning.org/cellpapers/cellpapers_v4n1.pdf

Trivette, C. M., Raab, M., & Dunst, C. J. (2014). Factors associated with Head Start staff participation in classroom-based professional development. *Journal of Education and Training Studies*, *2*(4), 32-45.

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Testing the Impact of Counseling over Time on Non-Urgent Undergraduate Life Satisfaction

Gary Blau[1], John DiMino[2], Iris Abreu[2] & Kayla LeLeux-LaBarge[2]

[1] Human Resource Management Department, Fox School of Business, Temple University, Philadelphia, USA

[2] Tuttleman Counseling Services, Temple University, Philadelphia, USA

Correspondence: Gary Blau, Human Resource Management Department, Temple University, 349 Alter Hall, 1810 Liacouras Walk, Philadelphia, PA, 19122, USA. E-mail: gblau@temple.edu

Received: May 25, 2017 Accepted: June 8, 2017 Online Published: July 4, 2017

doi:10.5539/jedp.v7n2p59 URL: http://doi.org/10.5539/jedp.v7n2p59

Abstract

Data for non-urgent undergraduate clients at a University Counseling Center (UCC) were collected using on-line surveys over three time-periods. Despite the expected diminishing number of participants over time, within-time correlations generally showed that level of mental distress and mental health concerns were negatively related to life satisfaction, while self-esteem was positively related. Using a smaller sample of matched-over-three time-periods clients, levels of mental distress and mental health concerns significantly declined, while life satisfaction significantly increased. The strongest changes for this complete data sample were found from Time 1 to Time 3 for reduced mental health concerns and increased life satisfaction. These findings reinforce that additional counseling sessions for undergraduates may be beneficial, and that spreading these sessions out may also be useful. Scientifically demonstrating to higher-level University administration that a UCC can help undergraduates in distress should ideally help the UCC to increase its allocation of university-based resources.

Keywords: life satisfaction, mental health concerns, impact of counseling, non-urgent undergraduates

1. Introduction

Choosing to seek counseling at a University Counseling Center (UCC) is an important decision in the lives of many undergraduates. The thought of sharing one's personal history and experiences to a stranger can be very scary (Rosenbaum & Liebert, 2015). Unfortunately, mental health issues continue to grow on college campuses, and there seems to be no abatement to the increased use of UCCs. Looking at data across 275 University Counseling Centers (Gallagher, 2014), 94% of UCC directors reported greater overall numbers of students seeking services, with 52% of UCC clients having severe psychological problems (up from 44% in 2013). The purpose of this study is to examine the impact of counseling services on life satisfaction and related correlates over time on a sample of undergraduates. These are important issues to consider as UCCs compete for university resources with other campus stakeholders (Hunt, Watkins, & Eisenberg, 2012).

1.1 Prior General Research on Undergraduate Life Satisfaction and Several Correlates

Life satisfaction among college students is a key component of subjective well-being (Diener, Suh, Lucas, & Smith, 1999). It is defined as an individual's evaluation of the quality of one's life against self-imposed standards (Diener, Emmons, Larsen, & Griffin, 1985). Recent studies have shown that life satisfaction is a protective factor against a non-suicidal self-injury in college students (Kress, Newgent, Whitlock, & Mease, 2015), and that life satisfaction mediates the relationships of attachment anxiety and avoidance to two domains of graduation-related concerns for college seniors: career-related and change and loss-related (Lane, 2016). Prior research has shown that self-esteem was positively correlated to life satisfaction, while neuroticism was inversely related to life satisfaction (Pavot & Diener, 1993). Diener and Seligman (2002) found that the happiest group of undergraduate students scored lower on several psychopathology scales (e.g., depression, psychopathic, family conflict) of the Minnesota Multiphasic Personality Inventory compared to less happy groups. Happiness was based on aggregated peer reports of affect balance, and self-reports of life satisfaction and affect balance collected over two months. This suggests that mental health concerns are negatively related to life satisfaction.

Using data from the National Comorbidity Survey Replication, Kessler et al. (2005) estimated that three-quarters of lifetime mental disorders (e.g., anxiety, mood, impulse and impulse control) have their first onset by the traditional college age range of 18-24. Using various pre-counseling and post-counseling measures, Brunner, Wallace, Reymann, Sellers and McCabe (2014, p. 264) called today's college students the "most stressed" generation of college students. When simultaneously comparing in-counseling versus not-in-counseling undergraduate samples, researchers have reported that students who are in-counseling have significantly lower academic adjustment, social adjustment and personal emotional adjustment (DeStefano, Mellott, & Petersen, 2001), and lower perceived well-being and mental health (Green, Lowry, & Kopta, 2003). This research collectively suggests that level of mental distress, mental health concerns and self-esteem are correlates of life satisfaction, and college students are an important sample to study.

1.2 Brief Nature of College Counseling

Given the increased demand of services on many university campuses (Gallagher, 2014), counseling with college students is typically of short duration. Using the data gathered from 1,698 college student clients across 42 universities, Draper, Jennings, Baron, Ozgur and Shankar (2002) found that the average number of sessions per client was 3.3. In another review, Ghetie (2007, p. 51) noted a median (midpoint of a frequency distribution) of 4-5 counseling sessions and a mode (most frequent number) between 1 and 2 sessions. Mahon et al. (2015) found that 37/124 (30%) of undergraduate clients completed a minimum of three counseling sessions, with the remaining 87 either never returning for a second session or dropping out after two sessions. While it can be very challenging to build a longer-term sample of undergraduates receiving counseling services, prior research suggests that even a short-term intervention can increase important undergraduate outcomes. Mahon et al. (2015) found a significant decrease in counselor-reported client symptom severity. Blau et al. (2016) found that the means of social connectedness and life satisfaction significantly increased over time for 28 undergraduates, with a median of four counseling sessions (after initial triage assessment). The current study measured a different set of antecedent variables, i.e., mental distress, mental health concerns, and self-esteem, as well as life satisfaction, and collected a longer-term sample of students in counseling (up to a median of 10 weeks, range of eight to 12 weeks), while testing two hypotheses. For the hypotheses below (H1 and H2), life satisfaction was defined as an individual's evaluation of the quality of one's life (Diener et al., 1985) and three different time periods were used to define "over time":

H1—level of mental distress and mental health concerns will be negatively related, while self-esteem will be positively related, to life satisfaction.

H2—over time counseling will reduce the level of mental distress and mental health concerns and increase self-esteem and life satisfaction.

2. Method

2.1 Samples and Procedures

Three undergraduate samples based on time length in counseling were collected in this study. All data were voluntarily collected at the main campus of a large state-supported urban university's UCC in the Mid-Atlantic region of the United States. The first sample is labeled as *initial counseling session sample or Time 1*. Both subsequent samples were derived from this initial sample and are labeled *Time 2* and *Time 3*. However, for any given client, the nature of their mental issue, as well as client/therapist availability did not allow for a standardized number of sessions or time periods to represent when clients filled out the *Time 2* or *Time 3* surveys. Given this non-standardized time-period and number of counseling sessions provided for participants before taking *Time 2* and *Time 3* surveys, a range of sessions and time-periods were used (reported below).

Data were collected for 261 undergraduates who registered for counseling services at the UCC between the fall of 2013 to the fall of 2016. Institutional Review Board (IRB) approval was given for all data collection, and data were collected using online surveys which went directly into Qualtrics' data bases. Lucas (2012) used a similar three-year period (2004 to 2007) to collect an undergraduate counseling sample. These undergraduate students had been designated at their initial counseling session (i.e., an in-person triage interview), as non-urgent (non-suicidal or risk to self/others) by the triage therapist. Students were asked if they wished to voluntarily participate in a research study looking at the impact of counseling on student outcomes. The *Time 1* clients took their survey after triage but before beginning counseling (i.e., intake). The time span between triage and intake could be anywhere from that day to over two to three weeks, depending on current client demand for services and therapist availability. Sample demographics for the Time 1 sample are reported below.

Undergraduate clients who filled out the initial survey (*Time 1*) were given the option of listing the last four digits of their nine-digit university identity number so that their responses, while receiving counseling, could be tracked over time but their individual identities would remain protected. Clients could take the *Time 1* survey either at the UCC's self-help center or at home using a survey link. Using their email addresses recorded at triage, clients were contacted for *Time 2* collection at four weeks from their triage appointment date and again for *Time 3* at eight weeks from date of triage appointment. A general email was sent out to all post triage clients on a rolling weekly basis during a semester. As a check, clients were asked what survey they were filling out (*Time 1*, *Time 2*, *Time 3*).

Of the 261 participants from *Time 1*, N = 70/261 (27%) filled out a second survey at *Time 2*. This *Time 2* sample reported a median of three counseling sessions after intake (range 1 to 8) over a period of six weeks (range 4 to 10 weeks). Of the *Time 1* clients, N = 35/261 (13%) filled out the third survey or *Time 3*. This *Time 3* sample reported a median of six counseling sessions after intake (range less than 5 to 12) over a period of ten weeks (range 8 to 12 weeks). Thus there is a client participation rate decline from N = 261 (*Time 1*) to N = 70 (*Time* 2) to N = 35 (*Time 3*). In order to improve the longitudinal response rate, a $50. random-drawing lottery for matched ID repeat-respondents was approved by the IRB part-way through the data collection process. However, this incentive had little impact in improving the matched respondent participation rate. This response-rate percentage decline is consistent with prior research (Lucas, 2012). In addition, archival data from the UCC study site showed that over the last three years, by 4 sessions (full intake plus three sessions after) over 60% of the non-urgent clients had stopped counseling. Across all three time periods there were only 20 clients who matched across time, when using their four-digit identification number. Despite assurances of confidentiality, clients may not have wanted their answers tracked longitudinally. This three-time match sample (N = 20) will be called *the complete sample*.

2.2 Measures

Respondent demographics. Demographics were collected only at Time 1. Data collected were: main mental health issue for presenting to counseling, gender, ethnic background, status entering university, current residential status, parents' highest education level, self-reported Grade Point Average (GPA), age, referral source for counseling, days waited for intake (after triage) and status as a full-time student (defined as taking over 12 credits/semester).

Level of mental distress, mental health concerns, self-esteem, and life satisfaction. These variables were measured at all three times. *Level of mental distress* was measured by asking clients to "rate your current level of mental distress that brought you to counseling". An eight-point Likert response scale was used, where: 1 = none, 2 = very little, 3 = a little, 4 = moderate, 5 = a lot, 6 = severe, 7 = very severe, 8 = unbearable. *Mental health concerns* were measured using six items. Items were "I have recently: (1) lost a lot of sleep over worrying, (2) felt constantly under some type of strain, (3) felt that I could not overcome my difficulties, (4) been feeling unhappy or depressed, (5) been losing confidence in myself and (6) been thinking of myself as a worthless person". These six items represent the Symptoms of Mental Disorder factor in the short-form of the General Health Questionnaire (Hu, Stewart-Brown, Twigg, & Weich, 2007). The 4-point response scale used by Hu et al. (2007, p. 1006) was expanded to a 7-point response scale to keep response scale consistency. The higher the total score the greater one's mental health concerns, with a Cronbach's alpha of .80 for the *Time* 1 sample. *Self-esteem* was measured using five items. Items were: (1) "I feel like I am a person of worth, at least on an equal basis with others"; (2) "I feel like I have a number of good qualities"; (3) "I am able to do things as well as most people"; (4) "I take a positive attitude towards myself"; and (5) "on the whole I am satisfied with myself". These are the positively worded items from the Rosenberg (1965) 10-item General Self-Esteem Scale. The other five reverse scored items from this scale were not used because research has shown that the negative items constitute a separate factor (Greenberger, Chen, Dmitrieva, & Farruggia, 2003). The higher the score the greater one's self-esteem, with a Cronbach's alpha of .92 for the *Time* 1 sample. *Life Satisfaction* was measured by using three items (Pavot & Diener, 1993). Items were: (1) "I am satisfied with my life", (2) "In most ways my life is close to ideal", and (3) "The conditions of my life are excellent". Data analyses across multiple samples has shown that these three items (of the five-item total) have the strongest item factor loadings and item-total correlations (Pavot & Diener, 1993). The higher the total score the greater one's life satisfaction, with a Cronbach's alpha of .86 for the *Time 1* sample.

2.3 Data Analyses

The first hypothesis (H1) was tested using correlation analyses on the four variables, i.e., level of mental distress, mental health concerns, self-esteem and life satisfaction, within each time-period for all complete data samples. The second hypothesis (H2) was tested using the General Linear Model (GLM) for each of the four variables. The GLM was appropriate for a repeated measures design (Stevens, 1996). There was no homogeneity of variance test with the GLM because only one group, i.e., the *complete sample*, was used. However, Mauchly's sphericity test, which is appropriate for 3 or more levels of a repeated measure factor (i.e., time), was calculated prior to testing between group differences (Stevens, 1996). A non-significant test indicated that the variances of the differences between all possible pairs of within-subject conditions (i.e., levels of independent variable or 3 different times) was equal and the subsequent multivariate F test was not inflated. For each GLM test, Mauchly's sphericity test was not significant, supporting the subsequent multivariate F test reported. If a significant multivariate F was found, the least squares difference post hoc test between groups was then used (Stevens, 1996). Since direction was specified in H2, an *a priori* one-tail test is justified (Stevens, 1996). Effect size for the largest significant difference between groups within a variable was also reported (Cohen, 1988).

3. Results

3.1 Demographic Comparison of Complete to Initial Sample

Table 1 shows a demographic comparison of the *complete sample* (N = 20) to the initial *Time 1* sample (N = 261). Overall, there was general consistency between the two samples, suggesting that the subsequent missing data after the *Time 1* sample was not due to any demographic variable measured, such as ethnicity, gender, or age (Roth, 1994). Comparing both samples to more general prior research: anxiety and depression were the top two mental health concerns of undergraduate clients (Barr, Krylowicz, Reetz, Mistler, & Rando, 2011), and males were generally less likely to seek counseling than females (Nam et al., 2010). Over a two-year period at a large urban community mental health center in the northwestern United States, Sparks, Daniels and Johnson (2003) found higher percentages of white (versus non-white) and self-referral (versus other referral) participants.

Table 1. Demographic variables for three-time complete versus initial samples

Variable	Complete (N = 20)	Initial (N = 261)
Main Mental Health Issue		
Anxiety	45%	42%
Depression	30%	28%
Gender		
Male	15%	25%
Female	80%	73%
Transgender	5%	2%
Ethnic background		
Non-white	30%	26%
White	70%	74%
Status		
No transfer	65%	69%
Transfer	35%	31%
Residential status		
Non-commuter	70%	76%
Commuter	30%	24%
Parents' education		
Less than 4-year college degree	25%	34%
At least 4-year college degree	75%	66%
Self-reported GPA	3.3	3.2

Age, M (range)	21 (18-29)	20 (18-30)
Referral Source		
Self	60%	59%
Other (e.g., family, friend)	40%	41%
Mean Days Waited for Intake	17	15
Full-time Student (12 plus credits)	95%	96%

3.2 Testing the Two Hypotheses

To test H1, i.e., level of mental distress and mental health concerns will be negatively related, while self-esteem will be positively related, to life satisfaction, Pearson correlational analyses were used. Within-time correlations for each complete sample data set are shown in Table 2.

Table 2. Means, standard deviations, and correlations among level of mental distress, mental health concerns, self esteem and life satisfaction within three separate times

Time 1 (N = 164)						
Variable	M	SD	1	2	3	4
1. Level of Mental Distress[a]	4.93	1.07	----			
2. Mental Health Concerns[b]	5.21	1.06	.47**			
3. Self-Esteem[b]	4.58	1.44	-.36**	-.50**		
4. Life Satisfaction[b]	3.76	1.52	-.43**	-.48**	.68**	----
Time 2 (N = 53)						
Variable	M	SD	1	2	3	4
1. Level of Mental Distress[a]	4.21	1.35	-----			
2. Mental Health Concerns[b]	4.72	1.29	.60**			
3. Self-Esteem[b]	4.76	1.61	-.45**	-.64**		
4. Life Satisfaction[b]	4.21	1.51	-.54**	-.60**	.75**	---
Time 3 (N = 35)						
Variable	M	SD	1	2	3	4
1. Level of Mental Distress[a]	3.83	1.42	-----			
2. Mental Health Concerns[b]	4.10	1.31	.60**			
3. Self-Esteem[b]	5.28	1.33	-.28	-.58**		
4. Life Satisfaction[b]	4.71	1.29	-.40*	-.63**	.62**	----

Note. * $p < .05$; ** $p < .01$ (two-tailed). Listwise deletion of data within each time.

[a]Level of Mental Distress, 1 = none, 2 = very little, 3 = a little, 4 = moderate, 5 = a lot, 6 = severe, 7 = very severe, 8 = unbearable.

[b]Mental Health Concerns, Self-Esteem, Life Satisfaction, 1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = neither agree nor disagree, 5 = somewhat agree, 6 = agree, 7 = strongly agree.

Using listwise deletion reduced the sample sizes from *Time 1* (N = 164) to *Time 2* (N = 53) and *Time 3* (N = 35). Despite the decrease in sample sizes over time, the results were generally consistent across each sample. Level of mental distress and mental health concerns exhibited a significant negative correlation to life satisfaction, while self-esteem showed a positive relationship to life satisfaction. In addition, the absolute value of the correlations

between level of mental distress, mental health concerns and self-esteem were .75 or less, indicating that these measures were distinct (Stevens, 1996). Overall, these results supported H1.

To test H2, i.e., over time counseling will reduce the level of mental distress and mental health concerns and increase self-esteem and life satisfaction, the GLM procedure was used. The results are shown in Table 3. Significant overall changes (multivariate F) in the *complete sample* were found for three of the four variables tested: level of mental distress ($p < .05$); mental health concerns ($p < .01$), and life satisfaction ($p < .05$). Level of mental distress declined, and the mean difference decline between *Time 1* (4.58) to *Time 3* (3.58) was significant. A large effect size of .79 (Cohen, 1988) was found for this decline. Mental health concerns also declined, and the mean difference declines between *Time 1* (5.05) and *Time 2* (4.65) to *Time 3* (3.94) were significant. For the largest difference in groups, the effect size was .93 (Cohen, 1988), which is classified as large. Life satisfaction increased and the differences between *Time 1* (3.80) and *Time 2* (4.15) to *Time 3* (4.82) were significant. A large effect size of .70 was found between the biggest difference groups. The multivariate F for self-esteem was not significant ($F = 2.36$). Although the descriptive data show smaller cell mean changes over time, self-esteem did increase. Overall, there was partial support for H2.

Table 3. General linear model testing significance of outcome variable, post hoc differences within subgroups, and largest effect size for complete sample

Outcome Variable	Level of Mental Distress			Mental Health Concerns			Self-Esteem			Life Satisfaction		
Multivariate F	F = 2.74*			F = 5.70**			F = 2.36			F = 4.58*		
Cell Means	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
	4.58[a]	4.05	3.58[b]	5.05[a]	4.65[a]	3.94[b]	4.56	4.79	5.30	3.80[a]	4.15[a]	4.82[b]
Effect Size[d]	*d* = .79			*d* = .93						*d* = .70		

Note. N = 20; T1 = Time 1; T2 = Time 2; T3 = Time 3.

Response scale for all outcome variables, 1 = strongly disagree to 7 = strongly agree.

[a,b]Within each variable, cell means that do not share the same superscript are significantly different at the 0.05 level by the Least Significant Difference post hoc test.

[d]Cohen's (1988) effect size, where *d* = M (NIC)-M (ICS)/√pooled SD for largest significant difference within outcome variable groups.

*p < .05; **p < .01 (both one-tailed).

4. Discussion

The research design of this study is rare in finding significant correlations among study variables over three time periods for an undergraduate sample of students receiving counseling services; as well as significant changes, with large effect sizes (Cohen, 1988), in three of four variables for the complete counseling sample. These results are consistent with a general Consumer Reports article on the benefits of psychotherapy (Seligman, 1995). The robustness of the findings, despite the loss in client participation over time, is encouraging. As level of mental distress and mental health concerns significantly decreased for clients, their life satisfaction significantly increased. These results are consistent with Lucas (2012) who found significant improvement in the "highly distressed" undergraduate client sub-sample between intake and their eighth session. Prior research (Kitzrow, 2009) has also shown that college students with higher perceived psychological distress were less likely to persist towards graduation, and depression was a significant predictor of both lower GPA and higher probability of dropping out (Eisenberg, Golberstein, & Hunt, 2009). The three scales used, mental health concerns, self-esteem, and life satisfaction, were shortened versions of longer scales. However, these shortened versions each demonstrated good scale reliability. Prior research working with college student samples suggests that shorter survey length can help with a higher response rate (Fan & Yan, 2010).

4.1 Study Limitations

Perhaps the biggest limitation of the current study is the loss in client sample size, especially from *Time 1* (N = 261) to *Time 2* (N = 70), but then also to *Time 3* (N = 35). Although such client loss is consistent with prior research (e.g., Draper et al., 2002; Ghetie, 2007; Lucas, 2012; Mahon et al., 2015), it is none-the-less discouraging. Natural client attrition due to treatment, i.e., clients sufficiently resolving their mental health issue

in a short time period, played a large role in these sample size decreases over time. The loss in sample size influenced the non-significant finding for self-esteem, which still did increase over time. One contributing factor for the loss between *Time 1* (after triage but before intake) to *Time 2* (median of three sessions after intake) may have been the wait time for intake (Lucas, 2012) which averaged over two weeks. Millennial college students have been stereotyped as "typically impatient, want-it-now" (Rickes, 2009). Prior research has shown that as wait time increased between triage and intake, clients were less likely to show up for the scheduled intake (DiMino & Blau, 2012).

In addition, when matching four-digit ID numbers, the complete data samples size over time periods was further reduced. Another contributing factor to difficulty in building a longitudinal sample is the "episodic" nature of counseling for many college students, which has been characterized as students seeking psychotherapy for a particular issue but then stepping away and taking a break from it, and returning later if needed (Webb & Widseth, 1988). This episodic nature is reflected in the uneven number of counseling sessions and time period ranges after *Time 1* for *Time 2* and *Time 3*. We do know that the *Time 2* sample had a median of three counseling sessions after intake, while the *Time 3* sample had a median of six counseling sessions. Unfortunately, no persistence/retention data were collected. Collecting such data may be important going forward. Prior research has shown that student satisfaction with their university was positively related to undergraduate retention (Schreiner & Nelson, 2013). It would be interesting to see if undergraduate life satisfaction also influences their retention. The final limitation was that only non-urgent clients were sampled.

4.2 Implications for Practicing College Clinicians and Resourcing UCCs

Scientifically demonstrating to higher-level University administration that a UCC can help undergraduates in distress should ideally help the UCC to increase its allocation of university-based resources (Bishop, 2010; Hunt et al., 2012). Linking student services to student life satisfaction can be important for improving the overall financial health of an institution. The current longitudinal findings about the positive impact of UCC services is an important part of that conversation. Such research efforts need to be continued as UCCs compete with other university stakeholders for resources (Castonguay, Locke, & Hayes, 2011). Aggregating common data across UCCs is one way to combat limited-sample-size at a particular UCC. Kopta et al. (2014) used a data set of 13,803 clients on the Behavioral Health Measure-20 (BMH-20) across 23 UCCs in the United States from 2006-11 to show that client treatment benefit on the BMH-20 peaked between 7 to 10 sessions.

In order to "hold" non-urgent triaged clients to increase their show-rates for their intake session, experimenting with such strategies, as a self-help center and group counseling may be helpful. Follow-up with no-shows to find out why they did not come to intake (DiMino & Blau, 2012) may also be helpful (e.g., used other coping resources, problem resolved, gave up waiting).

With regard to the study findings, the strongest change for the complete sample was found from *Time 1* to *Time 3* for reduced level of mental distress and mental health concerns, and increased life satisfaction.

The *Time 2* sample reported a median of three counseling sessions after intake (range 1 to 8) over a period of six weeks (range 4 to 10 weeks), while the *Time 3* sample reported a median of six counseling sessions after intake (range less than 5 to 12) over a period of ten weeks (range 8 to 12 weeks). This reinforces that additional counseling sessions for undergraduates may be beneficial (Lucas, 2012). In addition, spreading the sessions out may also be helpful. Surette and Shier (2017) recently found that while of number of sessions (Mean = 7) for their sample had a negative impact on somatization, depression and anxiety, having the sessions spread over a longer time period (M = 18 weeks) had a positive impact. Such research supports having the resources available in a UCC to avoid "time limited" treatment (Ghetie, 2007). Although many college students seem to successfully engage in short-term treatment (Mahon et al., 2015), knowing that additional sessions can lead to further benefits in terms of reduced mental health concerns and increased life satisfaction justifies a UCC having the resources in place, such as adequate staffing, to avoid an arbitrarily small number of sessions with college students (Lucas, 2012).

References

Barr, V., Krylowicz, B., Reetz, D., Mistler, B. J., & Rando, R. (2011). *The Association for University and College Counseling Center Directors annual survey*. Retrieved from http://files.cmcglobal.com/aucccd_directors_survey_monograph_2011.pdf

Bishop, J. (2010). The counseling center: An undervalued resource in recruitment, retention and risk management. *Journal of College Student Psychotherapy*, *24*, 248-260. https://doi.org/10.1080/87568225.2010.509219

Blau, G., DiMino, J., DeMaria, P. A. Jr., Beverly, C., Chessler, M., & Drennan, R. Jr. (2016). Social connectedness and life satisfaction: Comparing mean levels for two undergraduate samples and testing for improvement based on brief counseling. *Journal of American College Health*, *64*(8), 585-592. https://doi.org/10.1080/07448481.2016.1207645

Brunner, J. L., Wallace, D. L., Reymann, L. S., Sellers, J.-J., & McCabe, A. G. (2014). College counseling today: Contemporary students and how counseling centers meet their needs. *Journal of College Student Psychotherapy*, *28*, 257-324. https://doi.org/10.1080/87568225.2014.948770

Castonguay, L., Locke, B., & Hayes, J. (2011). The Center for Collegiate Mental Health: An example of a practice-research network in university counseling centers. *Journal of College Student Psychotherapy*, *25*(2), 105-119. https://doi.org/10.1080/87568225.2011.556929

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

DeStefano, T. J., Mellott, R. N., & Petersen, J. D. (2001). A preliminary assessment of the impact of counseling on student adjustment to college. *Journal of College Counseling*, *4*(2), 113-121. https://doi.org/10.1002/j.2161-1882.2001.tb00191.x

Diener, E., & Seligman, M. E. P. (2002). Very happy people. *Psychological Science*, *31*, 81-84. https://doi.org/10.1111/1467-9280.00415

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71-75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*, 276-302. https://doi.org/10.1037/0033-2909.125.2.276

DiMino, J., & Blau, G. (2012). The relationship between wait time after triage and show rate for intake in a non-urgent student population. *Journal of College Student Psychotherapy*, *26*(3), 241-247. https://doi.org/10.1080/87568225.2012.685857

Draper, M. R., Jennings, J., Baron, A., Erdur, O., & Shankar, L. (2002). Time-limited counseling outcome in a nationwide college counseling center sample. *Journal of College Counseling*, *5*(1), 26-38. https://doi.org/10.1002/j.2161-1882.2002.tb00204.x

Eisenberg, D., Golberstein, E., & Hunt, J. (2009). Mental health and academic success in college. *The B.E. Journal of Economic Analysis & Policy*, *9*(1), 1-37. https://doi.org/10.2202/1935-1682.2191

Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior*, *26*(2), 132-139. https://doi.org/10.1016/j.chb.2009.10.015

Gallagher, R. P. (2014). *National Survey of Counseling Center Directors*. Retrieved December 28, 2016, from http://www.iacsinc.org/

Ghetie, D. (2007). The debate over time-limited treatment in college counseling centers. *Journal of College Student Psychotherapy*, *22*(1), 41-61. https://doi.org/10.1300/J035v22n01_04

Green, J. L., Lowry, J. L., & Kopta, M. (2003). College students versus college counseling center clients. *Journal of College Student Psychotherapy*, *17*(4), 25-37. https://doi.org/10.1300/J035v17n04_05

Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem scale: Do they matter? *Personality and Individual Differences*, *35*, 1241-1254. https://doi.org/10.1016/S0191-8869(02)00331-8

Hu, Y., Stewart-Brown, S., Twigg, L., & Weich, S. (1997). Can the 12-item General Health Questionnaire be used to measure positive mental health? *Psychological Medicine*, *37*(7), 1005-1013. https://doi.org/10.1017/S0033291707009993

Hunt, J. B., Watkins, D., & Eisenberg, D. (2012). How do college campuses make decisions about allocating resources for student mental health? Findings from key participant interviews. *Journal of College Student Development, 53*(6), 850-856. https://doi.org/10.1353/csd.2012.0083

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593-602. https://doi.org/10.1001/archpsyc.62.6.593

Kitzrow, M. A. (2009). The mental health needs of today's college students: Challenges and recommendations. *NASPA Journal*, *46*(4), 646-660. https://doi.org/10.2202/1949-6605.5037

Kopta, S., Petrik, M., Saunders, S., Mond, M., Hirsch, G., Kadison, R., & Raymond, D. (2014). The utility of an efficient outcomes assessment system at university counseling centers. *Journal of College Student Psychotherapy*, *28*, 97-116. https://doi.org/10.1080/87568225.2014.883876

Kress, V. E., Newgent, R. A., Whitlock, J., & Mease, L. (2015). Spirituality/religiosity, life satisfaction, and life meaning as protective factors for nonsuicidal self-injury in college students. *Journal of College Counseling*, *18*(2), 160-174. https://doi.org/10.1002/jocc.12012

Lane, J. A. (2016). Attachment, well-being, and college senior concerns about the transition out of college. *Journal of College Counseling*, *19*(3), 231-245. https://doi.org/10.1002/jocc.12046

Lucas, M. (2012). Counseling on campus: Client persistence and progress. *Journal of College Student Psychotherapy*, *26*, 227-240. https://doi.org/10.1080/87568225.2012.685856

Mahon, M., Laux, J., McGuire Wise, S., Ritchie, M., Piazza, N., & Tiamiyu, M. (2015). Brief therapy at a university counseling center: Working alliance, readiness to change, and symptom severity. *Journal of College Counseling*, *18*, 223-243. https://doi.org/10.1002/jocc.12017

Nam, S., Chu, H., Lee, M., Lee, J., Kim, N., & Lee, S. (2010). A meta-analysis of gender differences in attitudes towards seeking professional psychological help. *Journal of American College Health*, *59*, 110-116. https://doi.org/10.1080/07448481.2010.483714

Pavot, W., & Diener, E. (1993). A review of the Satisfaction with Life scale. *Psychological Assessment*, *5*, 164-172. https://doi.org/10.1037/1040-3590.5.2.164

Rickes, P. (2009). Make way for Millenials!: How today's college students are shaping higher education space. *Planning for Higher Education*, *37*(2), 7-17.

Rosenbaum, P. J., & Liebert, H. (2015). Reframing the conversation on college student mental health. *Journal of College Student Psychotherapy*, *29*(3), 179-196. https://doi.org/10.1080/87568225.2015.1045780

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400876136

Roth, P. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, *47*(3), 537-560. https://doi.org/10.1111/j.1744-6570.1994.tb01736.x

Schreiner, L., & Nelson, D. (2013). The contribution of student satisfaction to persistence. *Journal of College Student Retention*, *15*(1), 73-111. https://doi.org/10.2190/CS.15.1.f

Seligman, M. E. (1995). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist*, *50*(12), 965-974. https://doi.org/10.1037/0003-066X.50.12.965

Sparks, W., Daniels, J., & Johnson, E. (2003). Relationship of referral source, race and wait time on pre-intake attrition. *Professional Psychology: Research and Practice*, *34*(5), 514-518. https://doi.org/10.1037/0735-7028.34.5.514

Stevens, J. (1996). *Applied Multivariate Statistics for the Social Sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Surette, T., & Shier, M. (2017). A common factors approach to supporting university students experiencing psychological distress. *Journal of College Student Psychotherapy*, *31*(2), 112-131. https://doi.org/10.1080/87568225.2016.1248233

Webb, R., & Widseth, J. (1988). Facilitating students' going into and stepping back from their inner worlds: Psychotherapy and the college student. *Journal of College Student Psychotherapy*, *3*(1), 5-15. https://doi.org/10.1300/J035v03n01_02

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Accelerated Cognitive Development—Piaget's Conservation Concept

Nobuki Watanabe[1]

[1] School of Education, Kwansei Gakuin University, Hyogo, Japan

Correspondence: Nobuki Watanabe, School of Education, Kwansei Gakuin University, Hyogo, Japan. E-mail: nobuki@kwansei.ac.jp

Received: July 6, 2017 Accepted: July 14, 2017 Online Published: August 3, 2017

doi:10.5539/jedp.v7n2p68 URL: http://doi.org/10.5539/jedp.v7n2p68

Abstract

Piaget's ideas have significantly influenced education and psychology, particularly the concept of conservation, which he had proposed as being acquired during the concrete operational stage. However, research conducted after Piaget found that children under the age of 6 are unable to understand his concept of conservation. However, more recent studies have found that three-year-olds may be able to acquire this concept, even when tested using the same tasks. But, this study addresses the issues of "fixity" and "reliability" for the concept of conservation. Then, the robustness (fixity and reliability) of Piaget's concept of conservation (numbers/length) was examined by observing a four-year-old child who demonstrated the possible acquisition of this concept at the age 3, in this study. It was found that the child was able to robustly maintain the concept. Therefore, the study shows the possibility of accelerated cognitive development for Piaget's concept of conservation. The reason may be that younger children have higher intelligence than those in previous generations. And, the grounds may be that of the influence of gene-environment interaction.

Keywords: accelerated cognitive development, conservation of numbers/length, gene-environment interaction, Jean Piaget

1. Introduction

Piaget's ideas have significantly influenced education and psychology, particularly the concept of conservation, which he had proposed as being acquired during the concrete operational stage. The conservation concept, which relates to the understanding of the equivalence of numbers, length, weight, mass, area, and volume requires abstract logical thought (transitive relations). The acquisition stage for numbers and length is generally determined using the following general Piagetian tasks and the stage has been clarified (Piaget, 1952; Ginsburg & Opper, 1969; Goswami, 1998; Siegler, DeLoache, & Eisenberg, 2003; Hetherington & Parke, 2003; Pastorino & Doyle-Portillo, 2013).

In the numbers task, two rows (R1, L1) with the same number of marbles are shown to children, who are then asked to show that the number of marbles in each row is the same. Subsequently, the distance between the marbles in one of the rows is increased or decreased, after which the children are asked if the modified row has the same number of marbles as the other row. After they answer, they are asked to show that the number of marbles in both rows is the same. The acquisition stage for this task is estimated to be 6–7 years old.

In the length task, two straight sticks of the same length are aligned in parallel (R1, L1). The children are then asked to show that the stick lengths are the same. Then, one stick is shifted from its parallel position, after which the children are asked again if the stick is the same length as the other. After the children answer, they again show whether the two sticks are of the same length. The acquisition stage for this task is also estimated to be 6–7 years. The conservation of mass has been found to occur around 7 years of age, the conservation of weight around 9 years of age, the conservation of area around 8–9 years of age, and the conservation of volume around 11 years of age.

Supplementary tests following Piaget's designs have demonstrated that acquiring the conservation of length and number before the age of 6 is rare (Goswami, 1998), with very little changes being found in the past 70 years, indicating that perception acquisition in children is relatively static. However, a "development acceleration phenomenon" has been observed in children's mental and physical development, for example, for cognitive development, it has been found that the ability of infants to read and write "hiragana" in Japanese has been

accelerating (Shimamura & Mikami, 1994); therefore, it is reasonable to surmise that there is a possibility of accelerated development in other specific areas.

The latest research suggests that even when the Piagetian design (tasks) has not changed, three-year-old children are also able to acquire the concept of conservation if the tasks are employed through play (daily life) (Watanabe, 2017).

However, this study addresses the issues of "fixity" and "reliability" for the conservation concept. As a child at a very young age might not fully remember the concept each time, it is necessary to continue conceptual testing as they age to verify conceptual fixity. In addition, because the correct answer for general Piagetian tasks is "the same" for all questions, when examining reliability, researchers have expressed a concern that children provide the correct answers in follow-up studies because of memorization or the "practice effect". Therefore, when conducting subsequent studies, it is necessary to devise a suitable method to demonstrate conservation acquisition robustness among children between the age of 4 and 6 years. If this can be identified, cognitive development acceleration for conservation could be concluded.

Therefore, this study investigated the fixity and reliability of the conservation concept in a four-year-old who had demonstrated the potential to acquire this concept at the age of three. Based on the results, this study suggests the possibility of accelerated cognitive development for the conservation concept.

2. Methods

To verify "fixation" and "reliability" for the conservation of numbers and length, the following procedure was enacted.

Research design: Single-case research methods were employed (Barlow, Nock, & Hersen, 2009) so as to continue the work from a previous study (there was also an issue regarding the physical difficulties of studying several people). The treatment variable (main variable) was the Piagetian task linked to play, including ordinary conversations. An A-B design was adopted because withdrawal procedures are naturally difficult.

Target: A four-year-old child (female) (from 4 years, 0 months to 4 years, 11 months), as in Watanabe (2017).

Characteristics of the subject child: She had not yet learned any special arithmetic. However, she had experienced mathematics in her ordinary conversations and during play in areas such as fractions, 3-D puzzles, a multiplication table song, and number and calculation quizzes. At 3, the child had been answering one item in the mathematics awareness survey (which has approximately 70 mathematical items) about once a month as part of her play (in a quiz or play format).

Stages by age in months: 4 years, 0 months was considered to be any day in the 1-month period from day zero of turning four until the day before turning 4 years, 1 month old. The study timing was arbitrary (cf., Watanabe, 2015).

Reliability Verification: This was the same as for Piagetian tasks; however, the child was not asked why the lengths and numbers were the same or different and, as a general rule, the final question was only asked once. All tasks were undertaken each month until the child was 4 years, 11 months old.

(1) Numbers

Five spheres were aligned in two rows (A1, A2) and the child was asked to show whether the number of spheres was the same in each row. Then, A2 was either lengthened or shortened (changed at least twice), after which the child was asked "which one is bigger" or "which one is larger" (Figure 1).

Figure 1. Numbers

(2) Length

・Two pens of the same length (A1, A2) were aligned and the child was asked to say whether they were the same length. Then, the position of A2 was offset or rotated up or down or left or right, after which the child was asked "which one is longer" (Figure 2).

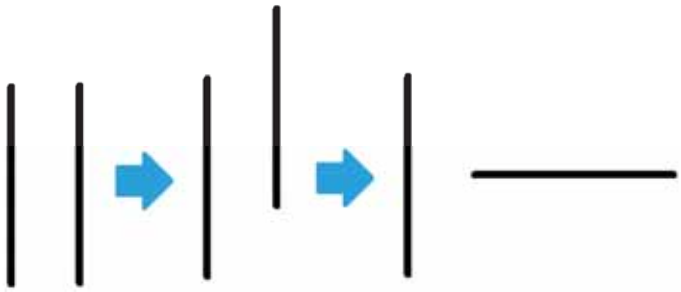


Figure 2. Length

Reliability Verification: The following changes were added to the Piagetian tasks (at 4 years, 4 months and 4 years, 10 months).

(1) Numbers

Five spheres were aligned in two rows (A1, A2) and the child was asked to show whether the number of spheres was the same in both rows. Then, A2 was lengthened or shortened (at least two times), after which the child was asked "which one is bigger" or "which one is larger". In the task progression, the number of spheres in one line was increased or decreased, after which the same question was asked again (Figure 3).

Figure 3. Numbers

(2) Length

Two pens of the same length (A1, A2) were aligned and the child was asked whether they were the same length. Then, A2's position was offset or rotated up or down or left or right and the child was asked "which one is longer?" In the task progression, the length of one pen was changed, after which the same question was asked again (Figure 4).

Figure 4. Length

3. Results

3.1 Fixity Results

Based on the results, the child attained a value of 1 (indicating a correct answer) at 3 years, 6 months of age for length and at 3 years, 7 months of age for numbers. Because this value was maintained at 1 for the subsequent 4–5 months, the concept seemed internalized until at least 4 years, 11 months (as it was clear, statistical analysis was not conducted). In other words, after the concept was acquired, it was maintained and internalized (Figures 5 and 6).

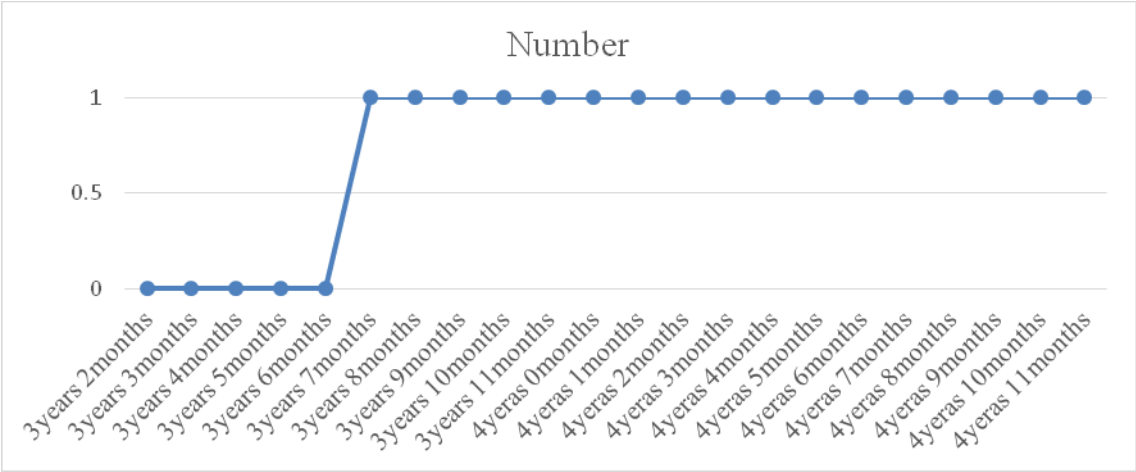


Figure 5. Numbers

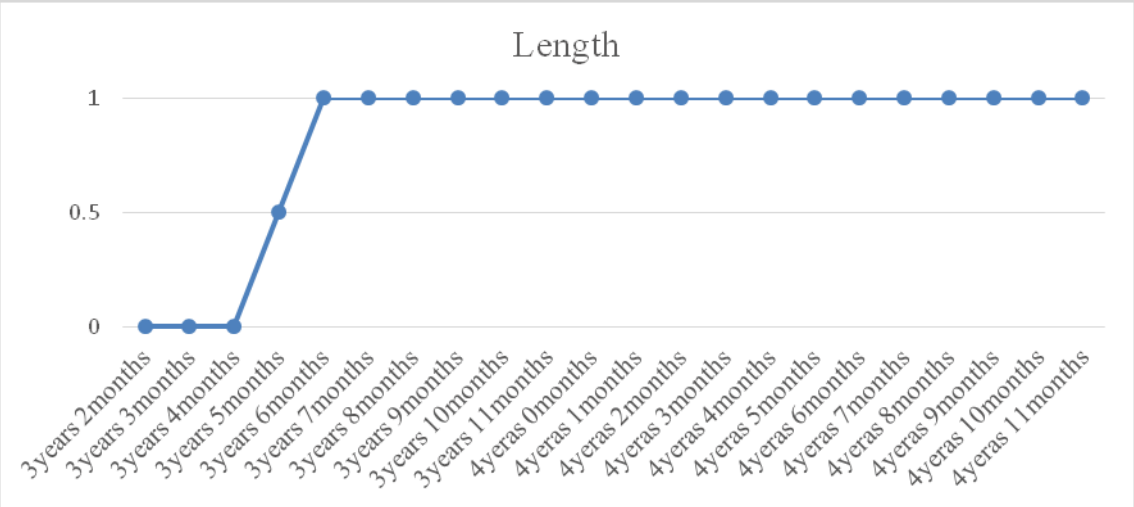


Figure 6. Length

3.2 Reliability Results

The child correctly answered all questions in the second survey for both length and numbers. Surveys were conducted once at age 4 years 4 months and again 6 months later at age 4 years, 10 months. In addition, a "size comparison" for which "the same" was not always the correct answer was incorporated in the middle of each of the surveys. Therefore, the possibility for the child to answer "the same" due to memorization was quite low; in other words, the responses indicated that the child had reliably acquired the conservation concept.

The above results showed that a four-year-old was able to maintain the conservation concept for both numbers and length.

4. Discussion

Prior research has found that children under the age of 6 are unable to understand Piaget's conservation concept (Goswami, 1998). However, more recent studies have found that three-year-olds may be able to acquire this concept, even when tested using the same tasks (Watanabe, 2017). Through regular surveys and scaffold testing, this study found that a four-year-old was able to robustly acquire the conservation concept for numbers and length, indicating acceleration in cognitive development from the previously presumed cut-off point of six years of age. Since the phenomenon was accepted by the investigation, we will discuss the theoretical factors of the possibility below.

The results from this study, therefore, serve to strengthen the proof for accelerated cognitive development for Piaget's conservation of numbers and length. It has been recently observed that younger children have higher intelligence than those in previous generations. For example, the finding of this study complies with Flynn's (2012) finding that the Intelligence Quotient (IQ) for children had increased. From around 1950 to 2000, both the Wechsler Intelligence Scale for Children (an intelligence test for 6-15-year-olds) and the Wechsler Adult Intelligence Scale (an intelligence test for 16-89-year-olds) showed a 15-point increase (Flynn, 2012). As Piaget's original research was conducted 70 years ago in 1940, if the IQ of younger children has increased since then, it is possible that the ability to acquire the conservation concept may have also shifted to a younger age. However, this study supports Flynn's (2013) analysis of the data on IQ tests and found that the IQ related to abstract problems had also increased significantly over the last century (Flynn, 2013). As Piagetian tasks generally assess logical thinking, it also seems valid that an increase in IQ could positively affect the ability to correctly solve the Piagetian tasks. This study agrees with Flynn's (2013) view that as the society has developed a more scientific attitude, there has been a commensurate increase in basic IQ, which he explained using "individual multipliers" and "social multipliers". Hence, aligning with Flynn's (2013) view, this study observes that individual multipliers alone, which are related to an individual's environment, are not sufficient to lead to an increase in IQ; however, because social multipliers, which are related to an individual's social environment, also affect analytical reasoning, these two together could have led to the increase in IQ with each generation. Overall,

however, there is little support by Flynn and other researchers for the postulation that environments alone affect IQ.

On this point, Heckman, after an extensive literature review, speculated that "gene-environment interactions may be central to explaining human and animal development" (Heckman, 2013, p. 16), an idea that was also supported by Rutter, a child psychiatrist, who claimed that as many psychological traits and mental disorders are multifactorial, this provides evidence that there are both genetic and environmental influences (Rutter, 2006). Behavioral geneticists, Asbury and Plomin also believe that people's environments operate in conjunction with genes, as genes are unable to activate without experience (Asbury & Plomin, 2014), which was in agreement with an earlier statement by behavioral geneticist, Spector, who claimed that "There are few if any examples of environmental factors without a genetic component, and conversely genes do not work alone and are usually dependent on the cells they live in and their environments. So, in a world where hundreds of genes are working together to influence a trait or disease, the old distinction between nature and nurture is simply no longer relevant" (Spector, 2012, p. 20).

Neuroscience studies have concluded that genes do not activate behavior; rather, they consist of a DNA sequence that contains all the relevant information necessary to produce a certain protein, and that the expression of the gene varies because of numerous factors, including environmental factors.

As research advances, the belief that there is a frontier between the innate and the acquired is disappearing, giving way to an understanding that there is interdependence between genetic and environmental factors in brain development (OECD, 2007). Medical scientist, Moalem, stated that advancements in scientific research have showed that genetic inheritance "can change and be changed by what we experience" (Moalem, 2014, p. 223). Further, one of the more important findings in 20th-century biology was the transfer of traits across generations and that the materialization of these traits in the complex development process involved interactions among genes as well as interactions between genes and the environment (Deutsch, 2012).

Epigenetics, a recent field of study, has been found to be the basis for these processes as environmental conditions may result in genes being turned off or on. While epigenetics does not consider certain phenomena, studies have shown that there is probably no separation between the impact of genetics and that of the environment (Asbury & Plomin, 2014; Rutter, 2006; Heckman, 2013; Spector, 2012), as exemplified by Shenk (2010) who claimed that "Everything we know about epigenetics so far fits perfectly with the dynamic system model of human ability. Genes do not dictate what we are to become, but instead are actors in a dynamic process" (p. 163).

Because of these new developments, it could be predicted that a child's phenotype significantly transforms when the three elements of genetics, individual environment, and (long-term) social environment are favorable. For future research into Piaget's conservation concepts and long-term social environments, it has been shown that the overall abstract IQ has significantly increased since Piaget's initial research. Further, a new and effective individual environment has been established, in which children have been "elevated to doing ordinary play with Piaget tasks" (Watanabe, 2017). Although gaining scientific evidence for changes in child genetics is quite difficult and should be a topic for future study, we know that the subject child's family lineage is prone to a phenotype that "excels at scientific subjects" because the father's maternal line has produced several science teachers. The child's father teaches university mathematics, her father's male cousin [mother's younger brother's child] teaches middle school mathematics, her father's other male cousin [mother's older brother's child] teaches middle school science, and her maternal grandfather is an elementary school teacher (specializing in science). In other words, the subject seems to have a genotype prone to this form of expression.

Considering all this information, it could safely be concluded that this study both practically and theoretically indicates the possibility of accelerated cognitive development related to Piaget's conservation concept. If the gene-environment interaction is at work, it is possible that individuals will have different abilities. Therefore, this single-case study seems to be an effective departure point for an exploration of the possibilities of acquisition of abilities.

As a first step in proving or validating strong hypotheses, single-case research methods can be effective; however, to validate the findings in this study, it is planned to increase the number of subjects in future research and to examine any acquisition changes in the other Piaget conservation concepts.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 16K01043.

References

Asbury, K., & Plomin, R. (2014). *G is for Genes: The impact of genetics on education and achievement*. London: Wiley-Blackwell.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston, MA: Allyn and Bacon.

Deutsch, J. (2012). *Le Ge`ne: Un concept en e´volution*. Paris: Seuil.

Flynn, J. R. (2012). *Are we getting smarter? Rising IQ in the Twenty-First Century*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139235679

Flynn, J. R. (2013). *Intelligence and human progress: The story of what was hidden in our genes*. London: Elsevier.

Ginsburg, H. P., & Opper, S. (1969). *Piaget's theory of intellectual development*. Eaglewood Cliffs, New Jersey: Prentice-Hall, Inc.

Goswami, U. (1998). *Cognition in children*. Hove, UK: Psychology Press, Taylor & Francis.

Heckman, J. J. (2013). *Giving kids a fair chance*. Cambridge, MA: MIT Press.

Hetherington, M. E., & Parke, R. D. (2003). *Child psychology: A contemporary viewpoint* (5th ed.). New York: McGraw-Hill.

Moalem, S. (2014). *Inheritance: How our genes change our lives and our lives change our genes*. New York: Grand Central Publishing.

OECD. (2007). *Understanding the brain: The Birth of a Learning Science*. Paris: OECD Publishing.

Pastorino, E., & Doyle-Portillo, S. (2013). *What Is Psychology? Essentials* (2nd ed.). Pacific Grove, CA: Wadsworth.

Piaget, J. (1952). *The child's conception of number*. London: Routledge & Kegan Paul.

Rutter, M. (2006). *Genes and behavior: Nature-nurture interplay explained*. London: Blackwell.

Shenk, D. (2010). *The genius in all of us: Everything you've been told about genetics, talent and IQ is wrong*. New York: Doubleday.

Shimamura, N., & Mikami, H. (1994). Acquisition of hiragana letters by pre-school children: In Comparison with the 1967 investigation of the National Language Research Institute. *The Japanese Journal of Educational Psychology*, *42*(1), 70-76. https://doi.org/10.5926/jjep1953.42.1_70

Siegler, R. S., DeLoache, J., & Eisenberg, N. (2003). *How children develop*. New York: Worth Publishers.

Spector, T. (2012). *Identically different: Why you can change your genes*. London: Weidenfeld & Nicolson.

Watanabe, N. (2015). A Basic Study on Development of School Mathematics Curriculum in Nursery School, Kindergarten, Center for Early Childhood Education and Care, and Home: Environmental setting for previous period of 3-year-old child. *Japan Journal of Mathematics Education and Related Fields*, *56*(1-2), 75-88.

Watanabe, N. (2017). Acquiring Piaget's conservation concept of numbers, lengths, and liquids as ordinary play. *Journal of Educational and Developmental Psychology*, *7*(1), 210-217. https://doi.org/10.5539/jedp.v7n1p210

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Theory of Mind Training in Children with Autism: Relating the Shared Attention Mechanism to the Theory of Mind Mechanism vs. Understanding Beliefs Training

Lama Bergstrand Othman[1] & Lana Collet-Klingenberg[1]

[1] Special Education, The University of Wisconsin-Whitewater, Whitewater, USA

Correspondence: Lama Bergstrand Othman, Special Education, The University of Wisconsin-Whitewater, 800 W Main Street, Whitewater, WI, 53190, USA. Tel: 1-262-472-3019. E-mail: Bergstrl@uww.edu

Received: July 4, 2017 Accepted: July 20, 2017 Online Published: August 3, 2017

doi:10.5539/jedp.v7n2p75 URL: http://doi.org/10.5539/jedp.v7n2p75

Abstract

The purpose of this training study was to examine two interventions that aim to improve the performance of students with Autism Spectrum Disorder (ASD) on False Belief Tasks (FBT) which examines the ability to recognize others' mental representation of the world. The first intervention involved drawing connections between the Shared Attention Mechanism (SAM) and the Theory of Mind (ToMM) Mechanism. The second intervention taught belief understanding and the fact that a person's beliefs about propositions may be false or true. To that end, we employed ABC and ACB multiple baselines across subject designs with matched controls. To assess generalization of learned skills, parents or teachers were interviewed. Results suggest that both interventions were effective.

Keywords: autism, Autism Spectrum Disorders—ASD, Theory of Mind—TOM, Shared Attention—SAM, intervention, social skills

1. Introduction

Theory of Mind (ToM) is the ability to infer other peoples' knowledge, desires, beliefs, and intentions (Bjorklund, 2005). The presence or absence of ToM can be investigated using False Belief Tasks (FBT) which examine one's ability to infer that someone else possesses a certain level of knowledge that may or may not be correct. First-order FBTs are related to the recognition of a person's knowledge, beliefs, or desires (e.g., one knows that Ali thinks it will rain today). Second-order FBTs are related to knowledge of a person's beliefs about another person's beliefs or mental state. For example, one knows that John believes that Ali thinks it will rain today (Baron-Cohen, 1996). This training study is based on the perspective that diminished competencies of ToM can be explained through a core deficit model; which entails a direct connection between ToM and its precursor skills such as shared attention mechanism-SAM (Fletcher-Watson, McConnell, Manola, & McConachie, 2014).

In a standard FBT, students watch a piece of candy being hidden in a drawer. Person X, who is present while the candy is hidden, leaves the room. While out, the candy is moved from the drawer to a new location. Students are then asked "when person X returns, where will he/she look for the candy?" (Wimmer & Perner, 1983). Most four-year old children successfully infer that person X will look for the candy in the drawer (Wellman et al., 2001). Another type of FBT involves showing students a distinctive box of candies that students are highly familiar with. Students are then asked what they think is in the box. The children say the name of the candy. The interventionist then reveals that there are no candies in the box but colors instead. Students are asked two questions: 1) what they thought was in the box prior to opening it; 2) what another person who didn't see the trick will think is in the box? Most four-year old children answer "candies" to both of these questions (Hogrefe et al., 1986). Regardless of the type of FBT, it has been established that children with autism have diminished competencies of ToM and hence fail FBTs (Bjorklund, 2005; Baron-Cohen, 1996; Perner et al., 1989; Leslie & Frith, 1988). Baron-Cohen (1996) proposed the "Agent-Attitude-Proposition" (for example, "Ali-thinks-it is raining"); In addition to the trueness of the proposition, typically developing children can comprehend that the

agent (in this case, Ali's) belief could be true or false. Most children with autism, on the other hand, comprehend the trueness of the proposition, but struggle with understanding the agent's belief.

ToM may be composed of four modules, each of which can work as an input to the other and can eventually engage ToM (Fletcher-Watson et al., 2014; Baron-Cohen, 1996). The Intentionality Detector (ID), which emerges at 0-9 months in typically developing children, is a simple module that exists in all animals with a nervous system. The ID is simply related to realizing or making inferences about the desire of an object or a living thing based on its movement. The Eye-Direction-Detector (EDD), emerging at 0-9 months in typically developing children, indicates that the child can detect an eye or an eye-like stimulus, recognize what the eye is looking at and based on that, infer that a person can see a certain object. Both the ID and EDD are dyadic representations of people or objects. The Shared Attention Mechanism (SAM), which emerges at 9-18 months in typically developing children, occurs when the child realizes that he/she and another person are looking at the same thing. The Theory of Mind module (ToMM), emerges between 18 and 48 months in normally developing children, and refers to a child's ability to represent other people's states of mind. SAM & ToMM are triadic mental representations. A deficit in shared attention is considered one of the early diagnostic criterion for ASD (DSM5; American Psychiatric Association, 2013). Children with autism are believed to have developed the ID & EDD but struggle with developing the SAM & ToMM (Baron-Cohen, 1996; Baron-Cohen, 1989). The implication of this discrepancy is that children with autism are unable to achieve a triadic type of representation or declarative comprehension (Broekhof et al., 2015; Kristen, 2011); one in which, a child and person X, for example, can attend to the same object/person or can share attention as social partners. Deficient triadic type of representation could explain why children with autism lack pretend and symbolic play, and consequently, they fail to relate to other people's different states of minds. However, this explanation was challenged by findings suggesting that many individuals with Asperger's syndrome succeed in FBTs despite their disengagement in pretend play and their sub-average understanding of others' mental states (Bowler, 1992). In the same vein, typically developing children use empathy and emotional routes to solve FBTs and similar social situations, whereas children with Asperger's syndrome who experience success in FBTs may be using alternative routes. In other words, their high cognitive abilities allow them to use cognitive routes that enable them to pass FBTs, and consequently, it takes them more time to solve FBTs than it takes typically developing children (Hermelin & O'connor, 1985) but they are still able to solve them.

Another proposed reason for the lack of ToM in people with of autism is a deficit in shared attention and ostensive behavior. That is, most children with autism demonstrate difficulty in imperative comprehension and in their ability to point to something or showing an object to someone (Broekhof et al., 2015; Kristen, 2011). Closely related to joint attention is ostensive communication, in which the child places a stimulus in someone's environment to direct that person's attention (Leslie & Happe', 1989). Students with autism lack social interest, which impedes their ability to develop shared interest (Lind & Bowler, 2009); hence, their inability to develop ToM could be the result of a lack of motivation rather than differences in their ability to represent other people's mental situation. The most popular explanation of the lack of ToM is related to Baron-Cohen's modules, which are ultimately related to a deficit in meta-representational abilities. The fact that the presence and absence of ToM is investigated using FBTs, which are based, for the most part, on language, suggests a connection between conversational skills and perspective taking (Ozonoff & Miller, 1995). Therefore, many training approaches seem to address specific social skills that are directly related to ToM. The results of such studies showed that individuals with autism demonstrated improvements in their overall performance on the FBTs; however, such improvement did not generalize to social skills, as rated by parents and teachers (Begeer et al., 2011; Golan & Baron-Cohen, 2006; Ozonoff & Miller, 1995). Many children use emotional routes to solve problems similar to those presented in FBTs, which suggests that a training approach that teaches emotions and belief understanding in addition to play-related components could be beneficial. However, some of these tasks showed no significant differences after such training (Handwin, Baron-Cohen, Howlin, & Hill, 1997). One prerequisite to understanding false belief is the ability to evaluate one's knowledge or lack of knowledge (Wellman & Liu, 2004; Wimmer & Gschaider, 2000). However, in many aspects of our lives, our knowledge is related to our visual perception or the see-know connection (Wimmer et al., 1988). This suggests that children with autism may have an interrupted perception-knowledge relationship, which is evident in their low performance on standard see-know tasks (Lind & Bowler, 2009; Baron-Cohen & Goodhart, 1993; Leslie & Frith, 1988; Perner et al., 1989).

Most research has used training skills that were not directly related to the hypothesis that SAM is an input of ToMM in the same manner that EDD is an input of SAM. Therefore, an intervention that directly accounts for

this type of relationship is worth investigating and is the purpose of this study. Pre- and post-intervention experimental designs are sound methods for examining the effect of a certain type of training on the ability to successfully complete FBTs. The time interval between the pre- and post-intervention evaluation should be moderately long to allow for sufficient training time but sufficiently short to ensure that the difference in performance is the result of the experimental factors. The purpose of this study is to examine two types of interventional program. One is based on Baron-Cohen's modules, which posit that EDD is an input of SAM-declarative attention. The other intervention program (described below) is training on false or correct beliefs that are based on visual perceptions.

2. Materials and Methods

2.1 Participants and Setting

Twelve students with autism, ages 9-12 years, participated in this study. Six students went through the training program, and the other six served as matched controls. The schools from which the students were selected were affiliated with School District One of the capital city of Amman, Jordan. Selection of the twelve students was based on four criteria: qualified for special education services for autism, verbal and language abilities that are at least at the 60^{th} percentile rank, and academic achievement at the 60^{th} percentile or higher. The fourth criterion was the inability to successfully complete FBTs. To determine whether the students met the criteria, the first author met with their special education teachers and school counselors and obtained parental consent to identify students whose records showed that they were recently given standardized tests of language and verbal abilities, an autism spectrum test, and a test of achievement. Students whose record did not include such standardized documentation or whose test(s) were administered three or more years ago were not considered for this study. Students who met the three testing criteria then took three similar FBTs. Those who failed at all three trials were considered candidates for the study (see Table 1). They were then matched according to age, gender, language and verbal abilities, and academic achievement. The final sample comprised 6 pairs/12 students.

Table 1. Demographics of participants and their matched pairs

	Age	Language and Verbal Abilities %ile	Most Recent Academic Achievement %ile	Eligibility for Autism	Gender
Pair1	9.11-10.2	60-63 %ile	62-65 %ile	√	Male
Pair2	10.6-10.8	79-81 %ile	63-65 %ile	√	Male
Pair3	10.11-11.0	68-69 %ile	67-70 %ile	√	Male
Pair4	11.3-11.7	65-66 %ile	63-67 %ile	√	Male
Pair5	12.4-12.4	64-66 %ile	68-70 %ile	√	Male
Pair6	12.7-12.9	69-70 %ile	73-78 %ile	√	Male

2.2 Procedures and Design

To investigate the presence or absence of ToM, FBTs were used. Each of these tasks could be repeated as many times as needed. The tasks were parallel to second-order FBTs, as described in the literature (Bjorklund, 2005; Baron-Cohen, 1996). Failure on FBTs was the final criterion for inclusion in this study, hence, these tasks were initially used to select participants. Later, the tasks were used in between interventions to investigate any improvement or acquisition of the ToMM. The implementation of each task did not exceed five minutes (see Tables 1).

Table 2. Example of False Belief Tasks for one session. Possible score 0-3

Session2-FBT1	Session2-FBT2	Session2-FBT3
The participant looks at a box of a popular packed treat (Tofeh). The interventionist opens the box and reveals that there is a pencil inside the box.	The participant checks a box of popular chocolate (Robert's). The interventionist opens then box and reveals that there are erasers instead.	The participant is looks at a kinder egg which is-popular in that area. The interventionist opens the egg and reveals that there is a paper clip instead of a little toy.
FB Question: "What did you think was in the box before opening it?"	FB Question: "What did you think was in the box before opening it?"	FB Question: "What did you think was in the egg before opening it?"

After obtaining permission from the participants' parents, the study was conducted at three schools within an eight-mile radius of one another and continued three-four times a week for six weeks. Graduate students who had a degree in special education and/or worked as special education teachers conducted the intervention program and the FBTs. Other adult volunteers assisted in conducting the study and observing the students' answers for later reliability checks. It is worth mentioning that because of the nature of the experiments, student volunteers who were siblings or friends of the participants were sought to help create the different experimental conditions in the interventions and the FBTs.

The first two FBTs (see Table 2) were only trials to help the participants become acquainted with the elements of the study. The three task implementations that followed were used as baseline data for each of the twelve participants. An ABC multiple baseline across-subjects design with matching controls was employed in this study. The baseline was A; the first training intervention (linking SAM to ToM) was B; and the belief-understanding training was intervention C. Intervention B aimed to teach students that we know what we see. This training is related to Baron-Cohen's concept of EDD and its relatedness to SAM and ToM. It follows the perspective of a core deficit model (Fletcher-Watson et al., 2014). Intervention C, in contrast, was based on training students on the idea that people may have true and false beliefs. Training C included activities aimed at understanding another person's perspectives. Each activity lasted 10 minutes. The training of each of the six students with autism took place in the library or in a room with one interventionist, another adult helper with whom the student was familiar, and a nondisabled sibling or a peer with whom the student with autism was familiar. For three of the students in the experimental group, the routine followed the ABC design for the first 3 weeks: 1) baseline (three sessions total); 2) intervention B was implemented (10 sessions total); 3) FBT was implemented (three tasks per session). For weeks 4-6, the routine was: 1) intervention C was implemented (10 sessions total); 2) FBT was implemented (three tasks per session); For the other three students in the experimental groups the routine followed the ACB design for the first three weeks: 1) baseline (three sessions total); 2) intervention C was implemented (10 sessions total); 3) FBT was implemented (three tasks per session). For weeks 4-6, the routine was: 1) intervention B was implemented (10 sessions total); 2) FBT was implemented (three tasks a session). The reason for using an alternate ABC/ACB design was to make sure that the type of interventions rather than the order of the interventions caused the change in the results.

To further clarify the procedures, the training of one student using the ABC design is described here: During a session of intervention B, an object with which the student was familiar with, a tennis ball, was hidden in a place while the student with autism and his peer/sibling were watching, and the interventionist asked the peer/sibling to leave the classroom. While the peer/sibling was outside the classroom and within sight of the student with autism, the interventionist removed the tennis ball and placed it in a different location, then asked the student with autism "Where do you think your (brother) will look for the ball when he comes back?" The interventionist and the adult helper coded the student's answer as 0 if he referred to the second location and 1 if he referred to the first location. During that period, a video camera recorded the trial; the examiner played the video from the beginning and prompted the student with autism in the following manner: "Did you see me when I changed the location of the ball? How did you see me? Who else saw me? What about your (brother); did he see me? Where was he?" Finally, the routine above was implemented again to understand the effects of the intervention on the student's answer to the FBT question "Where do you think your (brother) will look for the ball when he comes back?" Inter-rater reliability was determined afterwards using a category-to-category comparison. If the student could answer the FBT question 3 consecutive times, he exited the study without having to complete all 6 weeks of sessions.

Students in the ABC design who were unable to provide 3 or more consecutive correct answers on the FBT questions during intervention B were moved to intervention C after session 10, and students in the ACB design who were unable to maintain 3 or more consecutive correct answers to the FBT questions during intervention C were moved to intervention B after session 10. It is worth mentioning that in many research studies the FBTs and intervention close to intervention B of this study were conducted using toy characters and verbal or visual scenarios where students responded to the question by selecting a face or a toy character (Loukusa, Mäkinen, Kuusikko-Gauffin, Ebeling, & Moilanen, 2014; Fisher & Happé, 2005). In this intervention; however, the researchers attempted to create a relatively authentic environment where students with ASD, siblings/peers, and adults were actively involved in the experiment.

The training activities in intervention C were highly structured situations that emphasized the misrepresentations of others' knowledge and were saturated with teasers. One example of such training is as follows: the student with autism and his peer/sibling, together with the interventionist, are cutting papers to paste using a step-by-step visual organizer. The nondisabled peer was previously instructed to try to paste the paper before cutting it, which violated the orders in the visual organizer. After the students pasted the papers onto their cardboard, the interventionist asked the student with autism questions that helped him understand the peer/sibling's misconception of the activity procedures. The prompts were as follows: "Did he cut the papers? What do you think he didn't do right? Shall we go and tell him how it should be done? What shall we tell him? Can you show him the advanced organizer? Let's check and see if he will do it right this time." It is worth mentioning that each training activity in intervention C was selected according to the individual student's interest; hence, avoiding conflicts in desire among participants (Broekhof et al., 2015). For example, if a student liked cutting and pasting, the intervention was tailored around it. A student who liked to count and classify his little cars was given a counting and classifying activity to perform with his nondisabled peer/sibling in the fashion similar to that of the example above.

Engaging in a discussion about the mind, or about the emotional aspects of a story, has significantly helped typically developed students strengthen their concept of ToM (Bianco & Lecce, 2016). In intervention B, students with ASD were involved in a discussion that aimed to find the missing link "your brother didn't see, so he doesn't know". Likewise, in intervention C, students with ASD were involved in a three-way discussion to understand the inaccurate representation of the advanced organizer in the mind of the sibling/peer. Hence, this training study is highly structured around seemingly random conditions that aim to find the missing link between SAM and ToM.

3. Results

To compare the effectiveness of interventions B and C, performance of the three students in the AB design was compared to the performance of the three students in the AC design. Results showed that interventions B and C have similar degree of effectiveness (Figure 1). Regardless of whether they participated in intervention B or intervention C, the students in the intervention groups seem to have outperformed their matched controls (Figure 2). While the students in the experimental group participated in the study activities, the students in the control group went through the daily school routines that were provided by their schools according to each student's academic, behavioral, and emotional needs. This also applied to the students in the experimental group; that is, none of the special education services that they originally received at school were interrupted because of this study. The students in the control group went through the FBTs for the evaluation but did not participate in any of the intervention programs.

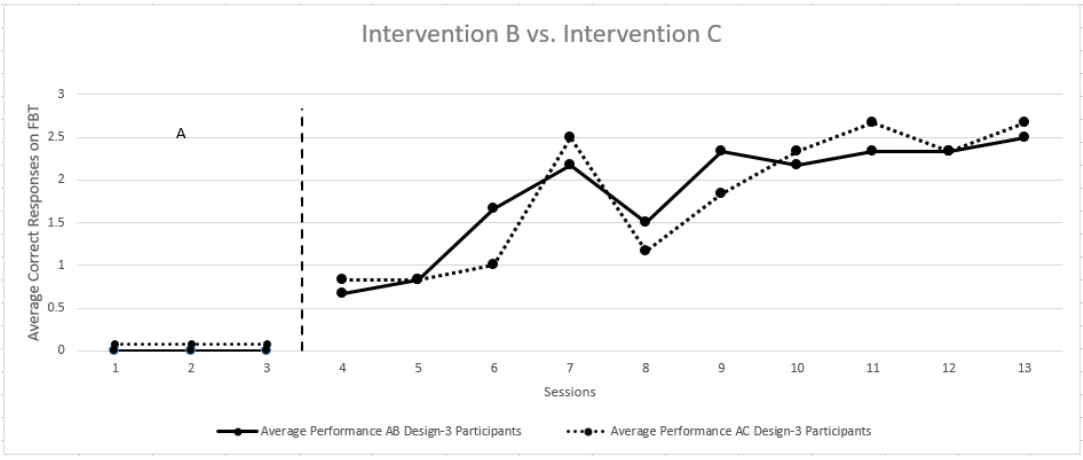


Figure 1. Connecting perceived information with knowledge, intervention (B) vs. belief understanding, intervention (C)

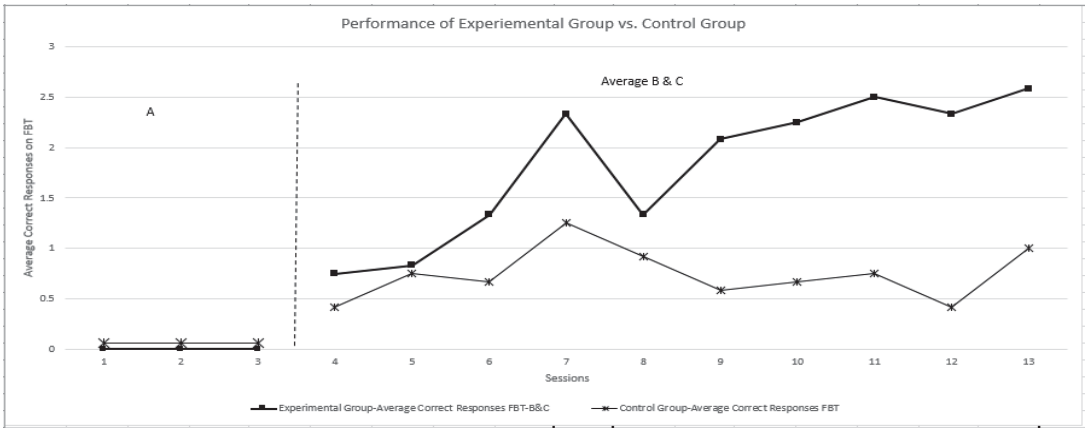


Figure 2. Comparison of performance on FBT between the experimental and control group

The student with autism who was assigned to the ABC design in pair one seemed to have benefited more from intervention C than from intervention B. Additionally, in both interventional programs, this student outperformed his/her control match (Figure 3). The second student in the ABC design seemed to benefit more from intervention B than the first student and exited the program early. Therefore, there was no need to apply intervention C (Figure 4). The third student, on the other hand, seemed to benefit more from intervention C than from intervention B. This student, like the first student, outperformed his/her matched control on both interventions (Figure 5).

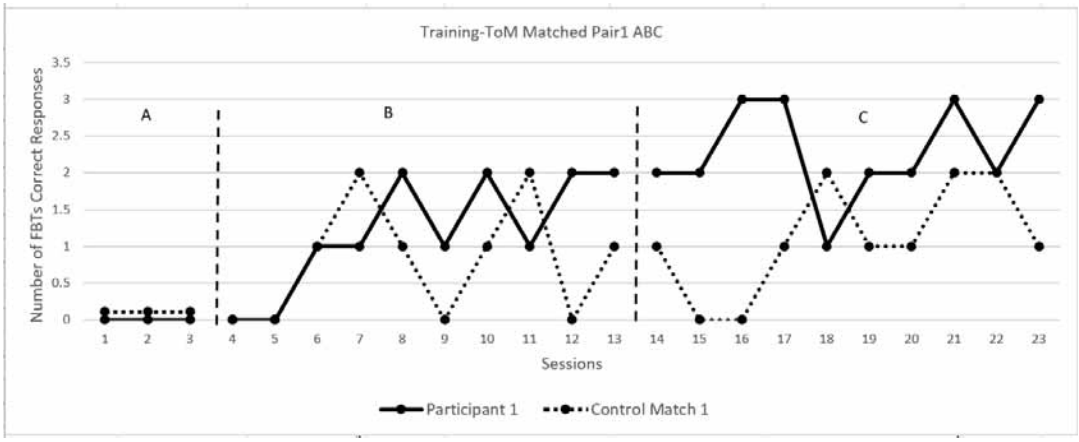


Figure 3. Matched pair one performance on FBT

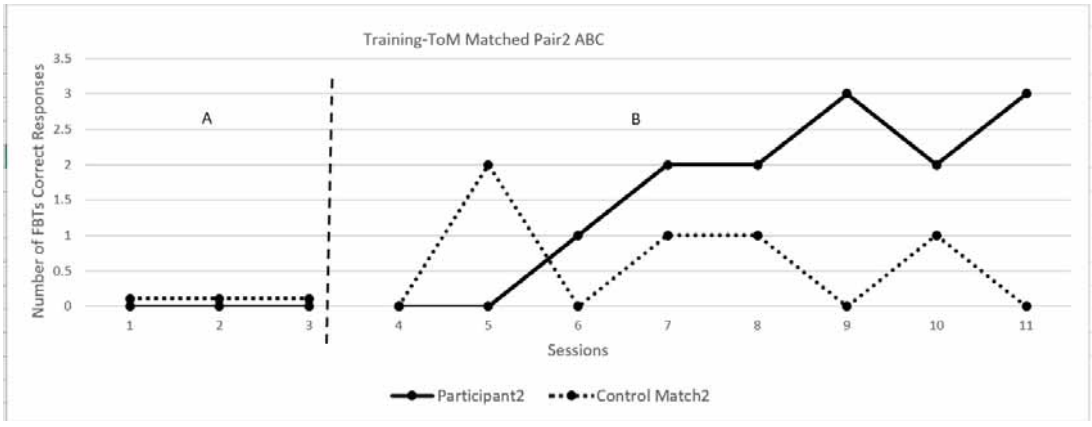


Figure 4. Matched pair two performance on FBT

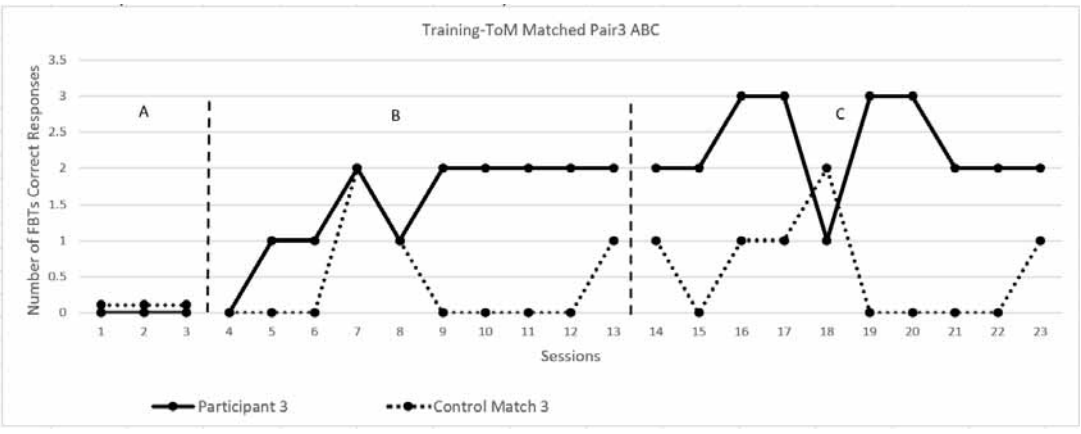


Figure 5. Matched pair three performance on FBT

The fourth student who was assigned to the ACB design (Figure 6) benefited more from intervention B than from intervention C and outperformed his/her control match on both interventional programs. The fifth student seemed to benefit more from intervention B than from intervention C and outperformed the matched control on

both interventional programs (Figure 7). Finally, the sixth student met the exit criteria early during intervention B and thus did not continue through intervention C (Figure 8).

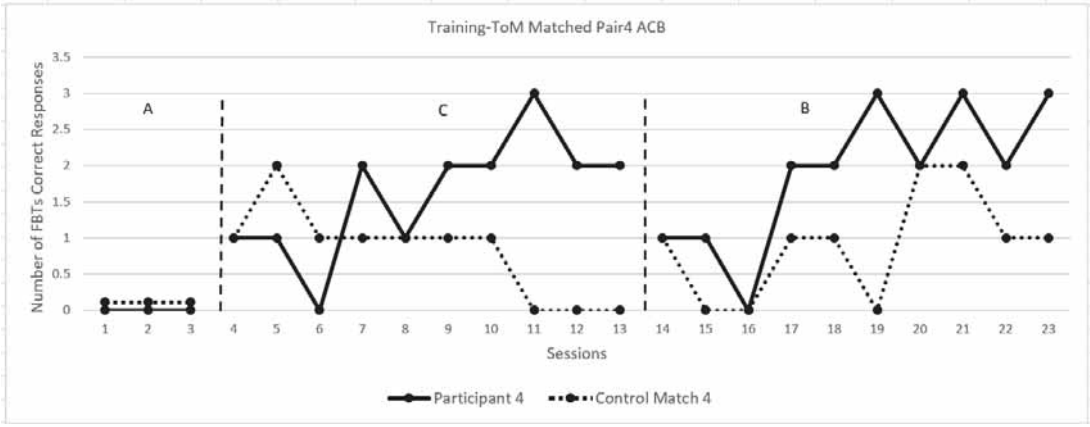


Figure 6. Matched pair four performance on FBT

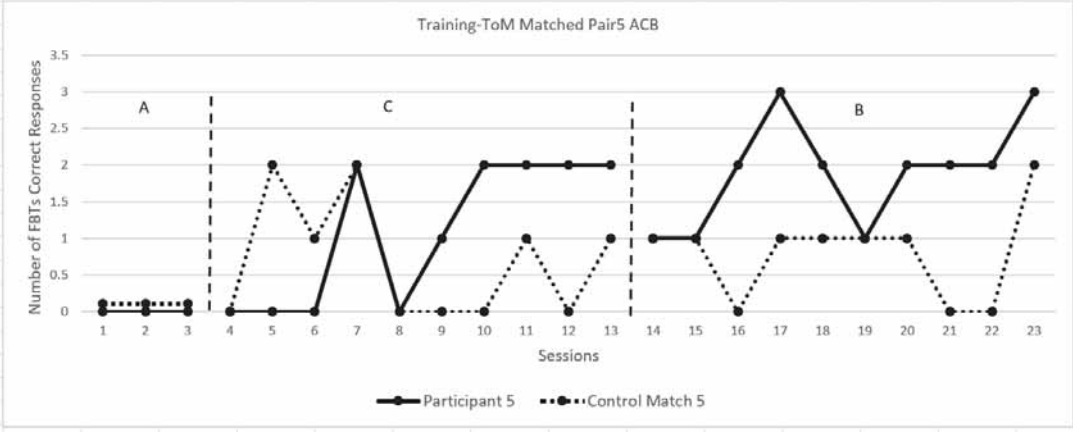


Figure 7. Matched pair five performance on FBT

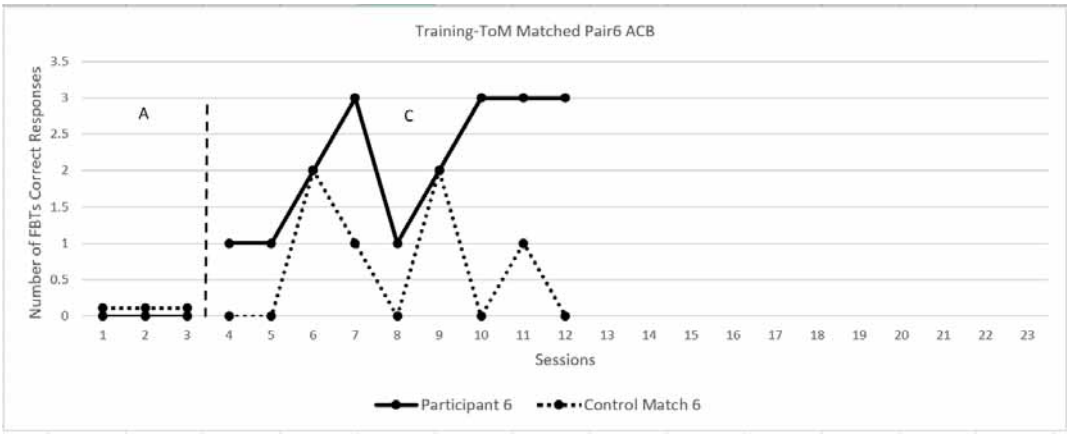


Figure 8. Matched pair six performance on FBT

After the intervention program concluded, in order to solicit any perceived changes in students' social communication, an informal interview was conducted with each student's parent or classroom teacher. The interview questions were based on ADEL Autism Spectrum Disorders Standardized Scale (Mohammad, 2002). For example, parents or teachers were asked if they noticed any differences in the students' ability to 1) pay attention to where someone is gesturing or pointing (imperative comprehension); 2) direct someone's attention to something by pointing or gesturing (declarative comprehension); 3) ask questions about future events; 4) play with others; 5) invite others to play with them; 6) move an object from one location to another in order to enhance its visibility; 7) ask questions such as "do you know?"; 8) rephrase for clarifications when asked; 9) provide justifications for his actions whether voluntarily or when asked; 10) provide justifications for other people's actions whether voluntarily or when asked. Despite the fact that there were differences among students in their level of engagement in one or more of the aforementioned areas, none showed any improvement in the level of engagement as perceived by their parents or teachers.

4. Discussion and Conclusion

Do students with autism fail ToM tasks because of information processing limitations or because of the lack of a genuine conceptual change in their cognitive abilities (Bjorklund, 2005)? In other words, is ToM a general domain competency or specific domain competency? Both perspectives acknowledge the need for a certain competence to develop ToM, but the former emphasizes the masking factors of early competence, whereas the latter acknowledges the presence of an evolutionary conceptual change that allows for the development of ToM. Students with autism lack the SAM and ToM modules, suggesting that the specific competency that allows transitioning from EDD to SAM is interrupted, and consequently, students with autism struggle to produce triadic representations of their social interactions. This single-case design research was an attempt to compensate for this interruption in a direct and systematic way. Moreover, the FBTs were expanded from their classical form by adding components that allowed training in addition to evaluation. For example, the use of video recording facilitated a shared attention simulation in which the interventionist and the student with autism were both viewing the same events in which the student had participated; thus, the events were not novel and were easy for the student with autism to relate to. Moreover, because of the interactive nature and constant reference to the events in the video recording, this intervention is expected to promote students' declarative comprehension or their ability to make a triadic mental representation (Broekhof et al., 2015; Kristen, 2011).

This simulation also allowed for natural reciprocity between the eye directionality detector (EDD; e.g., my brother was not in the classroom and thus did not see where the new location of the ball was), on the one hand, and between he did not see the new location and thus does not know, on the other hand. The types of prompts provided in intervention C also facilitated the student's ability to create a connection between a production of the other (peer/sibling) and how this other's misconception led to erroneous production. The other prompts in intervention C further facilitated the student's attempt to create a new understanding in the minds of the other (peer/siblings) to fix the mistakes that originated from the misunderstanding.

Reichow and Volknar (2010) classified research of evidence based instruction in social skills into three categories: 1) general social abilities (e.g., facial affect and ToM); 2) social interaction abilities (e.g., play skills and proximity); 3) social communication abilities (e.g., verbal social initiation, conversational skills, and shared attention). In order to capture the complexity of any one given social scenario, it was inevitable for the researchers to include different skills from each category of research classified by Reichow and Volknar (2010). While one could argue that it is seldom possible to draw conclusions about an experiment when the experimental factors are not purely isolated, others might argue that the nature of a given authentic social context doesn't naturally include one simple factor in isolation from another. After comparing results of the students in the two designs (ABC and ACB), it is readily apparent that more training leads to better performance on embedded trials. Whether the intervention started with training eye-knowledge relationship (B) and ended with belief training (C) or it started with belief training and ended with eye-knowledge training, the longer it lasted, the better the performance outcomes were. The FBTs are unique in regard to both their novelty and their linguistic and social elements; hence, further investigation into how these individuals were similar to each other and how they were different from the others would be helpful in understanding the effectiveness of the interventions. Upon review of participant demographics (see Table 1) the high performance of both participants two and six can be at least partially explained by their relatively superior language and verbal abilities. Future research should examine differences in language and verbal abilities among students with autism and how these differences may impact the efficacy of training programs such as the one incorporated in this study.

The results showed that compared with the students in the control group, those in the experimental groups improved in their conceptual ToM skills as demonstrated in embedded trials. Similar to Begeer et al. (2011) findings, parent and teacher reports did not support that students' generalized the learned skills outside of the specific intervention, which would suggest improved understanding of social behavioral. It was suggested by researchers that peer mediated intervention is an effective way to promote social behavior (Chan et al., 2009; Bellini, Peters, Benner, & Hopf, 2007). It was further suggested that parents' knowledge of the training program could expand the training condition to include a home setting; hence, increasing the chance of developing and generalizing social skills (Begeer et al., 2011). We recommend that this study be revised to include more effective roles of the peers and the parents of students with ASD. Likewise, it will be important for anyone implementing interventions such as those described here to include instruction in natural settings with structured opportunities to promote the student's ability to understand other people's points of view. Adopting a core deficit model to explain diminished competencies of ToM may require a longitudinal study that will examine ToM as a result of a chain of precursor skills (Fletcher-Watson et al., 2014). Unlike most ToM training programs that last a few weeks to a few months, we suggest a longitudinal training program that would allow for within-subject comparisons while cognitive, language, and social development take their natural course.

In summary, the interventions used in this single-case design research resulted in improved ToM skills related to False Believe Theory. Specifically, regardless of which intervention was implemented first, students in the intervention groups improved their ability to successfully navigate FBT tasks utilizing perspective-taking skills. Students who performed higher also had higher scores in verbal ability. Parent/teacher interviews did not suggest generalization of learned skills to natural settings. We suggest future training programs that include parents/care givers and peers over relatively long periods of time.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington: American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

Baron-Cohen, S. (1989). The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry, 30,* 285-297. https://doi.org/10.1111/j.1469-7610.1989.tb00241.x

Baron-Cohen, S. (1996). *Mindblindness: An essay on autism and theory of mind* (3rd ed.). Massachusetts Institute of Technology.

Baron-Cohen, S., & Goodhart, F. (1993). The "seeing leads to knowing" deficit in autism: The Pratt and Bryant probe. *British Journal of Developmental Psychology*, *12*, 397-402. https://doi.org/10.1111/j.2044-835X.1994.tb00642.x

Begeer, S. et al. (2011). Theory of mind training for children with autism: A randomized control trial. *Journal of Autism and Developmental Disorders, 41,* 997-1006. https://doi.org/10.1007/s10803-010-1121-9

Bellini, S., Peters, J. K., Benner, L., & Hopf, A. (2007). A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remedial and Special Education*, *28*(3), 153-162. https://doi.org/10.1177/07419325070280030401

Bianco, F., & Lecce, S. (2016). Translating child development research into practice: Can teachers foster children's theory of mind in primary school? *British Journal of Educational Psychology*, *86*(4), 592-605. https://doi.org/10.1111/bjep.12125

Bjorklund, D. F. (2005). *Children's thinking: Cognitive development and individual differences* (5th ed.). Belmont, CA: Wadsworth.

Bowler, D. M. (1992). Theory of mind in Asperger's syndrome. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *33*(5), 877-893. https://doi.org/10.1111/j.1469-7610.1992.tb01962.x

Broekhof, E., Ketelaar, L., Stockman, L., van Zijp, A., Bos, M. G. N., & Rieffe, C. (2015). The understanding of intentions, desires and beliefs in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*, 2035-2045. https://doi.org/10.1007/s10803-015-2363-3

Chan, J. M., Lang, R., Rispoli, M., O'Reilly, M., Sigafoos, J., & Cole, H. (2009). Use of peer-mediated interventions in the treatment of autism spectrum disorders: A systematic review. *Research in Autism Spectrum Disorders*, *3*(4), 876-889. https://doi.org/10.1016/j.rasd.2009.04.003

Fisher, N., & Happé, F. (2005). A training study of theory of mind and executive function in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *35*(6), 757-771. https://doi.org/10.1007/s10803-005-0022-9

Fletcher-Watson, S., McConnell, F., Manola, E., & McConachie, H. (2014). *Interventions based on the Theory of Mind cognitive model for Autism Spectrum Disorder (ASD)*. The Cochrane Library. https://doi.org/10.1002/14651858.CD008785.pub2

Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology*, *18*, 591-617. https://doi.org/10.1017/S0954579406060305

Hadwin, J., Baron-Cohen, S., Howlin, P., & Hill, K. (1997). Does teaching theory of mind have an effect on the ability to develop conversation in children with autism? *Journal of Autism & Developmental Disorders*, *27*(5), 519-537. https://doi.org/10.1023/A:1025826009731

Happe', F. (*1995*). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843-855. https://doi.org/10.2307/1131954

Hermelin, B., & O'Connor, N. (1985). Logico-affective states and non-verbal language. In E. Schopler, & G. Mesibov (Eds.), *Communication problems in autism*. New York: Plenum Press. https://doi.org/10.1007/978-1-4757-4806-2_15

Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, *57*, 567-582. https://doi.org/10.2307/1130337

Kristen, S., Sodian, B., Thoermer, C., & Perst, H. (2011). Infants' joint attention skills predict toddlers' emerging mental state language. *Developmental Psychology*, *47*(5), 1207-1219. https://doi.org/10.1037/a0024808

Leslie, A. M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology*, *6*, 315-324. https://doi.org/10.1111/j.2044-835X.1988.tb01104.x

Leslie, A. M., & Happe', F. (1989). Autism and ostensive communication: The relevance of metarepresentation. *Development & Psychopathology*, *1*(3), 205-212. https://doi.org/10.1017/S0954579400000407

Lind, S. E., & Bowler, D. M. (2009). Recognition memory, self-other source memory, and theory of-mind in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *39*(9), 1231-1239. https://doi.org/10.1007/s10803-009-0735-2

Lind, S. E., & Bowler, D. M. (2010). Impaired performance on see-know tasks amongst children with autism: Evidence of specific difficulties with theory of mind or domain general task factors? *Journal of Autism and Developmental Disorders*, *40*, 479-484. https://doi.org/10.1007/s10803-009-0889-y

Lukusa, S., Mäkinen, L., Kuusikko-Gauffin, S., Ebeling, H., & Moilanen, I. (2014). Theory of mind and emotion recognition skills in children with specific language impairment, autism spectrum disorder and typical development: Group differences and connection to knowledge of grammatical morphology, word-finding abilities and verbal working memory. *International Journal of Language & Communication Disorders*, *49*(4), 498-507. https://doi.org/10.1111/1460-6984.12091

Mohammad, A. (2002). *Adel ASD Standardized Scale*. Cairo: Al-rashad D. T. Press. Retrieved from https://slpemad.com/

Ozonoff, S., & Miller, J. N. (1995). Teaching theory of mind: A new approach to social skills training for individuals with autism. *Journal of Autism and Developmental Disorders*, *25*, 415-433. https://doi.org/10.1007/BF02179376

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. *Child Development*, *60*, 689-670. https://doi.org/10.2307/1130734

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, *40*(2), 149-166. https://doi.org/10.1007/s10803-009-0842-0

Robert, M. J., & Tager-Flusberg, H. (2004). The relationship of theory of mind and executive functions to symptom type and severity in children with autism. *Development and Psychopathology*, *16*, 137-155.

Vicker, B. (2003). Assessment day: Questions about the communication development of your young child with an Autism Spectrum Disorder. *The Reporter*, *8*(2), 18-21.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523-541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development*, *72*, 655-684. https://doi.org/10.1111/1467-8624.00304

Wimmer, H., & Gschaider, A. (2000). Children's understanding of belief: Why it is important to understand what happened. In P. Mitchell, & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 253-266). Hove: Psychology Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103-128. https://doi.org/10.1016/0010-0277(83)90004-5

Wimmer, H., Hogrefe, G. J., & Perner, J. (1988). Children's understanding of informational access as a source of knowledge. *Child Development*, *59*, 386-396. https://doi.org/10.2307/1130318

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Early Readers and Academic Success

Marie A. Leahy[1] & Nicole M. Fitzpatrick[1]

[1] Education Department, Chestnut Hill College, Philadelphia, Pennsylvania, USA

Correspondence: Marie A. Leahy, Education Department, Chestnut Hill College, Philadelphia, Pennsylvania, 19118, USA. Tel: 1-215-242-7757. E-mail: leahym@chc.edu

Received: July 19, 2017 Accepted: July 28, 2017 Online Published: August 22, 2017

doi:10.5539/jedp.v7n2p87 URL: http://doi.org/10.5539/jedp.v7n2p87

Abstract

This study examined the degree to which the age that a child learns to read affects his or her future academic success. In recent years, society and school districts, as well as an increasing number of parents, have been urging children to read at earlier ages. Therefore, in order to investigate the research question, an extensive survey was developed and electronically dispersed to individuals eighteen years of age or older. The survey was completed by 220 respondents, collected, and evaluated. The survey results provided quantitative data on respondents' demographic backgrounds as well as their childhood reading and academic histories. A significant number of respondents, 85%, said they regarded the age a child learns to read as important. Over 82% of early readers, ages three and four, described their overall academic success as either "Excellent" or "Very Good", which was 12% higher than the respondents who did not learn to read until age six or older. The factors that were not impacted by the age the respondents learned to read was whether or not they went to college or earned advanced degrees. Therefore, while academic success is often perceived as getting good grades in school, is that really enough? The respondents who learned to read at an early age generally earned good grades in school, but that did not necessarily translate to being more likely to go to college or earn an advanced degree, which is a strong measure of overall academic success.

Keywords: reading, early education, academic success, college and career ready, Common Core Standards

1. Introduction

Societal expectations, countless conflicting research findings, and competitive spirits have been some of the factors that have contributed to parents drawing significant attention and importance to the age at which their children learn to read. While some parents have tried to teach their children to read prior to formal schooling, others have allowed their children to learn at their own pace or at the age that most children are taught to read, in early elementary school. But a common goal that most parents share is that they want their children to be the best, brightest, and most successful students possible. According to Guernsey (2011), television advertisements for baby flashcards, an increase in the accessibility of the internet, countless parenting books, and the growing presence and weight of standardized tests have been some of the more recent causes for the increased pressures that parents and schools have put on children to learn to read. Since reading is one of the most important and valuable skills that every child needs to acquire and consistently improve, there is much controversy and differing opinions about how beneficial and effective early reading is to a student's eventual reading comprehension skills and overall future academic success (Kern & Friedman, 2014).

Although some research studies promote teaching children to read as early as possible, others have found that there are a variety of factors and variables that influence not only a child's future academic success, but also their readiness to learn to read. According to Kern and Friedman (2008), such factors have been socio-emotional maturity, parent involvement, personality, and gender. Lately, teachers have been put in a compromising position when it comes to this issue since they try to bridge the gap between parent expectations and what is educationally and developmentally appropriate for their students. While most families, educators, and educational standards are aligned to enrich children's minds and to prepare them for their academic futures, there are some philosophical differences that the three segments hold when it comes to children and the best process in which their long-term academic achievement is fostered.

Assessing the appropriate age to teach children to read and how valuable a factor it is to student academic success has major implications not only for students, but also for teachers, families, school curriculum, standardized tests, and Common Core Standards. Therefore, a literature review on past research studies and relevant articles has been completed and discussed within the body of this research paper. A survey has been developed, distributed, and analyzed to gain insight on the reading experiences and academic successes that people from a broad range of backgrounds have had in order to compare reading age and future academic success with other variables. So, the question that has been researched, examined, and evaluated is: To what degree does the age at which a child learns to read affect his or her academic success?

2. Literature Review

2.1 Introduction

Numerous studies have found that reading is one of the most valuable fundamental skills that children learn growing up, which aids them in their academic success at future grade levels. With all of the research on the benefits of reading, there has been a debate about the best age to teach children to read, causing two differing philosophical perspectives. The first is that children are never too young to learn how to read. These proponents have cited great amounts of research indicating that children who learn to read prior to their formal schooling benefit more academically than those who learn to read later in childhood or during the primary grades. The other side of the debate focuses on the developmental readiness of children to read, and does not see a great benefit in attempting to teach babies and toddlers to read as early preparation for formal schooling (Caron & Ponder, 2014). Parents, educators, and children are those who are most affected by this debate since parents want their children to be prepared and to be successful in school, while teachers recognize the developmental concerns and administrative pressures regarding the age at which children learn to read, and what age will prove to be most effective. Children are at the hands of research and what their parents want, which sometimes conflicts with what educators are trained to teach, based on standards, teaching strategies, and developmental considerations.

2.2 The Importance of Reading

Learning to read is the skill that a child will use for the rest of his or her life, which will ultimately help and affect their future academic success. Reading has been strongly linked to academic achievement, since reading is incorporated into every subject at higher grade levels (Caron & Ponder, 2014). Therefore children who are able to read and who have more advanced vocabularies and reading comprehension skills tend not only to do better in school but also to score higher on standardized tests. Children who struggle to learn how to read, or who give up and do not learn at all, have more difficulties with the rest of their education, or unfortunately, are forced to drop out. This happens since the student might be too embarrassed to continue to attend formal schooling (Reading to Young Children, 2012).

Children develop literacy skills and an awareness of language long before they are able to read. These language skills are linked to academic success since language development is fundamental to all areas of learning. Children who lack a strong foundation of language awareness and literacy skills early in life are more likely to fall behind in school leading to less academic success (Reading to Young Children, 2012). Based on prior research from the Institute of Education, children who read for pleasure are likely to do significantly better in school than their peers. Not only is learning to read important, but developing a love for reading is important when it comes to continued academic success (Battye & Rainsberry, 2013).

There is much debate about whether it is best to teach children to read early at younger ages or even prior to the start of their formal education, or if it is better to teach children to read during first and second grades when they are developmentally prepared. Prior research has shown that children learn the fastest in their first six years of life, which may bolster the philosophy that teaching children to read as early as possible is best (Benefits of Early Reading, 2010). Research has also shown that children must be developmentally and psychologically prepared and ready, otherwise learning to read may be more of a challenge, or even just a superficial skill, which the child does not know how to nurture or advance if uninterested (Kern & Friedman, 2014). Since reading is one of the most cognitively challenging skills for a child to learn, it is important that children be taught at the proper age so that there is less of a struggle and more ease based on prior language, vocabulary, and sound acquisition.

2.3 Advantages of Learning to Read Early

Children who are bright, curious, and developmentally ready to learn how to read are able to start at an early age. Some learn at home prior to the start of their formal education, while others learn in preschool or kindergarten.

According to Caron and Ponder (2014), parents who answer their children's questions, engage in dialogue, and read to their children, are the ones who set their children up to learn to read at earlier ages. This is because whether parents realize it or not, they are helping to build the foundation that every child needs in order to learn how to read. By putting language and words together, and familiarizing themselves with how books are held, the way words are written, and the sounds that different letters make alone or when joined with vowels or other letters, children begin to make connections and gain confidence with their understanding and abilities.

Early literacy experts argue that children at ages four and five should be exposed to some building blocks of reading. For example, they should learn the letters of the alphabet and their sounds, so that those components of reading are readily "retrievable" from memory when children are asked to start reading words in later grades (Guernsey, 2011). Whether children actually learn to read at age three or age seven, it is just as important for children to have gotten the language, alphabet, and sound exposure, which will aid in their reading development and advancement. Since research has found that children learn the fastest in the first six years of life, many proponents of teaching children to read at earlier ages defend their position by explaining how reading can be picked up faster and can be used throughout their early years to gain more information through the books they read and gain exposure to new vocabulary, moral lessons, and other subject areas (Benefits of Early Reading, 2010).

Growing children's vocabularies has been shown to bolster children's academic success as well as to improve standardized test scores. This is another reason why so many academics as of late have repositioned their views from supporting children's first reading age from six or seven to earlier ages. Another reason why the recommended reading age has dropped recently is because of the decline in literacy rates and the push for lower school entrance ages (Caron & Ponder, 2014). To help combat the growing decline in literacy, many feel that children who learn to read at younger ages will develop stronger reading foundations and abilities, which will set them up for success in their futures.

2.4 Advantages of Learning to Read Later in Childhood

Traditionally, experts have believed that children should learn to read in their first years of schooling, first and second grades. In the past, early childhood education was not as readily available or as common, but lately pre-kindergarten and kindergarten programs have sprouted, which has caused an increase in child enrollment into these classes (Caron & Ponder, 2014). This growing trend has added not only to the debate about what is the best age to teach children to read, but also what is the best age to begin formal schooling for children? Experts agree that children have individual learning speeds and styles, and mature at their own rate, along with other developmental, social, and biological variables that affect their readiness to begin formal schooling and learning to read. All children have their own interests and personalities, which translate to individual temperaments, all of which affect the age at which a child is ready and able to learn to read (Kern & Friedman, 2014).

Researchers agree that there is a limit to how young you can go when teaching a child to learn to read, and that it makes no sense to try to teach a baby or toddler how to read, when developmentally speaking, they are just not ready to begin. Although there are parents and educational advocates, like Janet Doman of the Institute for the Achievement of Human Potential, who try to use reading flashcards with babies and toddlers, there are many other parents and educators in the field who believe that children should read when they are ready but should always be exposed to books and be read to by their parents and families. The National Early Literacy Panel published a report in 2009, which was later criticized by some experts due to its narrow view of scientific studies, which de-emphasized the importance of play and conversation in language development, a critical foundation to success in reading. Yet in that report, there was an important synthesis of research on what types of early experiences are most closely related with successful reading, including exposure to letters and an awareness of phonemes in pre-kindergarten and kindergarten (Guernsey, 2011).

There are decades of peer-reviewed research that focuses on the need for a "balanced approach" that is committed to helping children to communicate using spoken and written language, while also helping them to identify individual letters through the alphabet, and to recognize the use of print by holding and observing books (Guernsey, 2011). A more recent study conducted by Kern and Friedman (2014), aimed to gather follow-up data from the Terman Life Cycle Study to explore how age at first reading and age at school entry affect grade school academic achievement, along with some other long term educational and psychological adjustments. This study found that early reading was associated with early academic success, but less "lifelong educational attainment" and worse "midlife adjustment". While Kern and Friedman's research discussed other studies' findings about early reading abilities and how they are directly and indirectly related to long-term reading success, this journal

article explored how precocious reading ability can dissipate over time, or a child may not be psychosocially prepared or developed enough to enjoy the early reading benefits, and therefore may have a more difficult time adjusting through the academic grade levels. This study's findings support the philosophy that exposing children to books and encouraging good reading behaviors and activities are beneficial, but they will not have all of the positive long-term educational successes if they are not psychosocially developed or ready to read or advance in grade level.

Learning to read later in childhood also allows the student to be helped and taught by a professional educator. These educators have been formally trained and can provide phonics instruction, reading techniques, word recognition activities, and model the reading process to their students. Children who learn to read during their first formal years of education can be monitored by a professional, which allows the teacher to facilitate, scaffold, correct, and even intervene during the reading process when appropriate. Children who learn to read later in their childhood may also benefit by having different types of books at their disposal in the classroom and school library, which a child who learns earlier in their childhood may not have access to. Having an early elementary teacher encourage reading and provide opportunities to read during the school day also encourages positive literacy behaviors and habits, which promotes recreational reading and instills a love and passion for literature (Cullinan, 2000).

2.5 Variables Affecting Reading and Academic Success

There are a number of variables that play a role and impact not only the age at which children learn to read, but also their future academic success. There are environmental factors, biological factors, and behavioral factors that can make a major difference in how children learn, and how advanced or behind they might become in later grades. Some of these variables include the following: parental level of education, socioeconomic level, race and culture, family make up and siblings, gender, personality, parental involvement, age at which child begins school, teacher opinions and biases, and technology (Kern & Friedman, 2014). The highest level of education that a parent has completed does have an impact on children and their academic success. If a parent dropped out of school, then they may not place an important value on education, or may feel insecure about their own knowledge and reading level competencies. This scenario is more likely to have a negative academic effect on a child, since parental encouragement and involvement is so important during the critical years that children learn how to read. On the other hand, if a parent has completed higher education, then they may hold education in higher regard and as a top priority, being more involved with their child's literacy development and academic achievement (Ortiz, Folsom, Otaiba, Greulich, Thomas-Tate, & Connor, 2012).

The socioeconomic background that a child comes from is another variable that can impact their reading age and future academic success. Children who come from poverty may not receive the proper nourishment to develop physically, emotionally, psychologically, or intellectually in the same way that children from higher social classes frequently do. When a child's basic developmental needs are not met or fostered, then learning to read is typically delayed, and formal schooling will be more of a challenge for that student. Environmental factors that are related to socioeconomic status and affect a child's reading age and academic success include living environment and neighborhood. If a child is living in a dangerous area with violence occurring at all hours of the night, and is sharing a bedroom with five other siblings, then they have some additional concerns than does the child who grows up in a safe middle class suburb, who is put to bed in their own room every night, at a reasonable time, with no outside disturbances (Ortiz et al., 2012). Race and culture can have a significant influence on children's reading age and academic success, since some cultures believe that education is up to schools and teachers, and is not a responsibility of the parent or family. Children whose parents do not assist with homework or educational lessons prior to the start of their formal schooling are at a disadvantage, since that extra reinforcement has been proven to bolster a child's education. Unfortunately, minority children statistically come from lower socioeconomic backgrounds than do Caucasian children, which is part of a larger issue that negatively affects those children and their future academic success (Lesaux, 2012).

Family make up is a variable that is two-fold. In single parent homes a child may not get as much parental attention or guidance as one in a two-parent home receives. A single parent may need to work double shifts to provide for the household, resulting in daycare or limited quality time. A two parent family may be able to afford to have one parent stay home with the child during their early years and upbringing, which likely provides more opportunities for the child to be read to by the parent (Ortiz et al., 2012). Battye and Rainsberry (2013) discussed the findings that having older siblings had a negative effect on children's test scores in three subject areas (math, spelling, and vocabulary) but particularly for vocabulary. Children who have older siblings may not get the same

opportunities to speak and ask questions as only children or oldest children do, and they may also be spoken for, or may not need to speak to get what they want, need, or want to know. According to Reading to Young Children (2012), if a child does have an older sibling, it is important to encourage the older sibling to read to the child since they will receive individual attention, and will learn to model their older brother or sister.

Other variables such as gender and personality can have an effect on children's academic success and reading age, since males and females develop and mature at different rates. Every child has a different personality and varying preferences in terms of what to do, how to do things, what interests them, and how they learn most effectively (Kern & Friedman, 2014). A parent's level of involvement with their child, and their education, is also an important factor that contributes to a child's academic success and reading age, since children learn from their parents, and engage in conversation, ask questions, and read together. When parents are involved with their child's education they become aware of educational milestones that are expected, and typically try to facilitate in any way possible at home. The age at which a child enters formal schooling, whether that be pre-kindergarten, kindergarten, or right into first grade will also have a role in a child's academic success. Not only do children benefit from the socialization opportunities that formal schooling offers, but they also have the opportunity to learn about different subjects, engage in read alouds, and explore learning stations in the classroom. Being exposed to books, reading centers, the alphabet, and spoken language earlier in life does help children learn to read when they are ready. Entering a grade when a child is not prepared, or developmentally mature enough, can cause academic struggles, and can make the reading process difficult and unpleasant for the student (Ortiz et al., 2012).

2.6 Conclusion

Based on the research and findings discussed, there is still a great deal of work, and studies to be completed, in order to provide answers as to how much of an impact the age at which children learn to read truly relates to their academic success. There will always be a differing of opinion in terms of what parents want for their children, and whether they will urge their children to read as young as possible, or whether it is their philosophy to let their children read and develop at their own pace and with the help of their school teacher. However, there is growing concern that the increase in testing at earlier ages will influence the recommended reading age, pushing the age at which children begin to read lower and lower, which will have negative consequences for our students' future academic success. On the other hand, based on previous research, it is now understood that exposing children to language, the alphabet, books, and new vocabulary as early as possible does benefit children's eventual reading skills. This in turn strengthens their academic success, since reading and reading comprehension carry over into other academic subjects. Being taught to read, or just naturally learning to read at age four or age six, can be equally beneficial in terms of a student's academic success if the proper instruction, materials, books, and developmental timing is respected, encouraged, and available. The goal is for children to develop a love and passion for reading, while improving and exercising the skill at every grade level.

3. Methods

A survey of nine demographic questions and twenty questions on reading and academic history was designed based on previous research and the literature review. In the summer of 2016, the survey questions were submitted to the Institutional Review Board at Chestnut Hill College and granted approval. Each survey question had a specific purpose, either to provide specific demographic information on the respondent or to help explain the respondents' reading habits and academic history. Ultimately, the goal of the survey was to compare demographic backgrounds, the age the respondent learned to read, and general academic experience, while examining the degree to which the age a child learned to read affected his or her academic success, taking all research variables into consideration. The demographic questions that were asked were a way of developing a history for each respondent. Since there are many variables that can have a role and impact a person's academic success, it was important to understand each respondent's background and how his or her experiences compared to others. Nomination sampling was utilized during the fall of 2016. Four surveys were initially sent electronically. The first four respondents forwarded the survey and then the pattern continued. All participants were at least eighteen years of age and lived in the United States. The format of the survey was intended to be appealing and user friendly in hopes of garnering a large return. Two hundred twenty-three responses were collected and three responses were removed because more than half of the questions were not answered. The quantitative data obtained from the remaining 220 surveys was analyzed using Excel.

4. Results and Discussion

Of the 220 responses, 76% were female and 24% were male with the age breakdown being as follows.

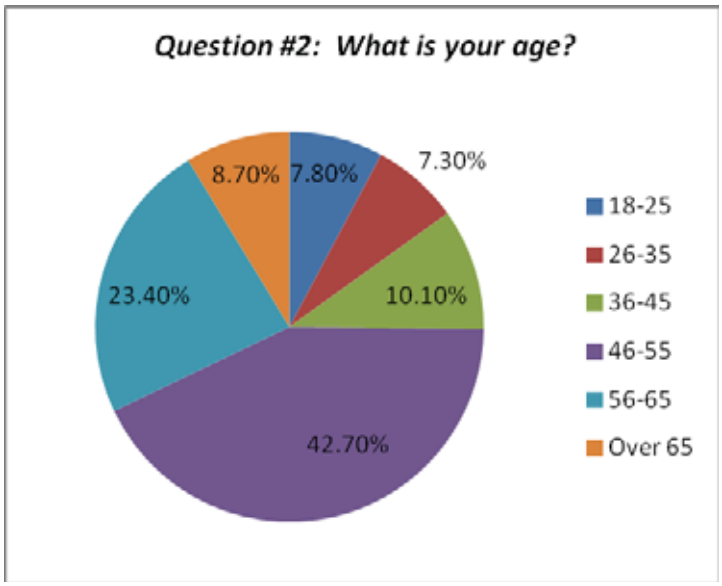


Figure 1. Question #2 on Survey: What is your age?

The age at which each respondent learned to read was varied. Two respondents did not answer question #11 (To the best of your knowledge, at what age did you learn to read?). Based on this data, the age the respondents learned to read was not significantly affected by socio-economic status, the level of education of either parent, the year they were born, or whether they came from a single-parent or two-parent home. The extent to which race was a factor could not be determined from this data set because 91.8% of the respondents were white/Caucasian.

Table 1. Age child learned to read

Age	***# of respondents***	***% of 218*** *(Two respondents left question blank.)*
3	19	8.71
4	43	19.72
5	66	30.27
6	73	33.48
7	15	6.88
8	1	0.45
10 or over	1	0.45

As indicated in the research, males tend to mature later than females and thus often learn to read later than females. Since about three-fourths of the number of respondents were female compared to about one-fourth being male, the age at which respondents of each sex learned to read cannot be compared in this study. It is interesting that the two respondents who learned to read at age eight and ten or over were both male.

In answer to question #26 (Do you think the age you learned to read is important?), 67% of the respondents replied, “Yes, definitely,” 18% replied, “Somewhat important,” and 15% replied, “No, not really.” Clearly, a

significant number of respondents, 85%, regarded the age they learned to read as either extremely important or somewhat important.

The respondents who self-reported that they viewed their overall academic success as either “Excellent” or “Very Good” was as follows:

Table 2. Respondents who viewed their overall academic success as either “Excellent” or “Very Good”

Respondents who learned to read at age 3 or 4	51/62	82.25%
Respondents who learned to read at age 6 or older	63/90	70%
Respondents who learned to read at age 7 or older	9/17	52.94%

The respondents who learned to read at an early age, three or four years old, and described their overall academic success as either “Excellent” or “Very Good” was 82.25%, which was significantly higher than the respondents who learned to read at age six or older who described their overall academic success as either “Excellent” or “Very Good”. Based on the self-determination of the participants in this study, the age at which a child learns to read significantly affects his or her overall academic achievement.

However, getting good grades in school can be very subjective, and what one person values as “good” may be considered “excellent” by someone else. In addition, getting good grades is not the sole determinant of academic success. In addition to grades on a report card, successfully going on to higher education is another determinant of overall academic achievement. The percentage of respondents who learned to read at the age of three and went on to attend at least “some college” was 94.73, which is noteworthy. Surprisingly, the percentage of respondents who learned to read at the age of four and went on to attend at least “some college” was significantly lower at 83.72, which was almost identical to the percentage of respondents who learned to read at the age of six, which was 83.56%. Interestingly, the percentage of respondents who learned to read at the age of five and went on to attend “some college” was 92.42, only slightly lower than those who learned to read at the age of three.

Table 3. Respondents who attended at least “some college”

Respondents who learned to read at age 3	18/19	94.73%
Respondents who learned to read at age 4	36/43	83.72%
Respondents who learned to read at age 5	61/66	92.42%
Respondents who learned to read at age 6	61/73	83.56%

As the respondents progressed through their academic careers, the age they learned to read did not seem to predict the level of education they obtained. The respondents who learned to read at the age of seven or older and who earned a Master’s Degree was 35.29% compared with those who learned to read and the age of three or four at 29%.

Table 4. Respondents who obtained a master’s degree

Respondents who learned to read at age 3 or 4	18/62	29%
Respondents who learned to read at age 6 or older	29/90	32.22%
Respondents who learned to read at age 7 or older	6/17	35.29%

The respondents who earned a doctoral degree learned to read at ages four, five, and six with a range of 4.54%-6.97%. Even though 6.9 % of those who learned to read at age four earned a doctorate compared with 4.54% of those aged 5 and 5.48% of those ages six, the differences are insignificant. The one fact that may be worth mentioning is that although 35.29 % of respondents who learned to read at age seven or older earned a master’s degree, none of the respondents who learned to read and age seven or higher earned a doctoral degree.

Table 5. Respondents who obtained a doctoral degree

Respondents who learned to read at age 4	3/43	6.97%
Respondents who learned to read at age 5	3/66	4.54%
Respondents who learned to read at age 6	4/73	5.48%

5. Conclusion and Recommendations

A child who learned to read at the age of three or four is considered an early reader. The majority of people, as well as the majority of respondents in this study, learned to read in kindergarten and first grade at the age of five or six. Over the past two decades there has been a growing trend to get children reading before they enter kindergarten. With the Common Core Standards children in kindergarten are expected to read and be assessed on large amounts of texts, often causing frustration on the part of the children, teachers, and parents. As a result, a new trend is emerging where parents are delaying the start of kindergarten for their children and waiting until they are six years old to enroll them rather than at the typical age of five so their children will not be "behind" and will do well on their tests. History has shown that the vast majority of children will learn to read by the end of first grade because that is the natural time in their growth and development where reading typically takes effect. Pushing children to read before they are ready is truly an injustice. As a result, an inordinate amount of young children now need extra support in the forms of tutoring, additional help from classroom aides, extra small group work, extra homework, and summer school. If so many children need so much extra support to learn to read, they are not being taught at the proper level. Learning to read does not need to be such a chore. It will happen because that is the natural progression once reading readiness skills have been acquired. There will always be some children who will naturally begin reading at an early age, but the vast majority of children learn to read at age five or six.

Although random sampling was not utilized in this study and the results cannot be generalized, the findings shed light on the ongoing conversation of the importance of reading at an early age and the impact on overall academic success. Clearly, learning to read at an early age can give children a boost of confidence and help foster a positive attitude toward school, but it is not the sole determinant of whether or not children will have overall academic success. Often times the understanding of academic success is getting high test scores and good grades, but that should not be the litmus test. Overall academic success should take into consideration whether or not a person was able to go onto higher education and become a productive member of society. The amount of pressure to begin reading at an early age that is put on children today by parents, teachers, schools, and school districts seems unnecessary. In this age of high stakes testing, children begin being tested on their reading skills and comprehension as soon as they are first registered for kindergarten. If the goal of the Common Core Standards is to get children "college and career ready", the age at which they learn to read is not a critical factor.

References

Battye, C., & Rainsberry, M. (2013, September 11). *Reading for pleasure puts children ahead in the classroom, study finds*. Retrieved from http://www.ioe.ac.uk/89938.html

Benefits of Early Reading. (2010, January 1). Retrieved September 7, 2014, from http://www.teachreadingearly.com/benefits-of-early-reading.php

Caron, C., & Ponder, C. (2014). *What's the Best Age to Teach Your Child to Read*? Retrieved September 17, 2014, from http://www.learn-to-read-prince-george.com/best-age-read.html

Cullinan, B. (2000). Independent reading and school achievement. *School Library Media Research*, *3*(1523-4320), 1-24. Retrieved from http://www.ala.org/aasl/sites/ala.org.aasl/files/content/aaslpubsandjournals/slr/vol3/SLMR_IndependentReading_V3.pdf

Guernsey, L. (2011, May 12). *Learning to Read: How Young is Too Young*? Retrieved September 17, 2014, from http://www.huffingtonpost.com/lisa-guernsey/learning-to-read-how-youn_b_860964.html

Kern, M., & Friedman, H. (2014). Early educational milestones as predictors of lifelong academic achievement, midlife adjustment, and longevity. *Journal of Applied Developmental Psychology*, 419-430. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2713445/

Lesaux, N. (2012). Reading and reading instruction for children from low-income and non-english-speaking households. *The Future of Children*, *22*, 73-88. https://doi.org/10.1353/foc.2012.0010

Ortiz, M., Folsom, J., Otaiba, S., Greulich, L., Thomas-Tate, S., & Connor, C. (2012). The componential model of reading: Predicting first grade reading performance of culturally diverse students from ecological, psychological, and cognitive factors assessed at kindergarten entry. *Journal of Learning Disabilities*, 406-417. https://doi.org/10.1177/0022219411431242

Reading to Young Children. (2012, January 1). Retrieved September 7, 2014, from http://www.childtrends.org/?indicators=reading-to-young-children

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Counseling over Time as a Correlate of Non-Urgent Undergraduate Institutional Commitment

Gary Blau[1], John DiMino[2], Iris Abreu[2] & Kayla LeLeux-LaBarge[2]

[1] Human Resource Management Department, Fox School of Business, Temple University, Philadelphia, USA

[2] Tuttleman Counseling Services, Temple University, Philadelphia, USA

Correspondence: Gary Blau, Human Resource Management Department, Temple University, 349 Alter Hall, 1810 Liacouras Walk, Philadelphia, PA, 19122, USA. E-mail: gblau@temple.edu

Received: August 9, 2017 Accepted: August 22, 2017 Online Published: September 20, 2017

doi:10.5539/jedp.v7n2p96 URL: http://doi.org/10.5539/jedp.v7n2p96

Abstract

The general purpose of this study was to examine counseling services as a correlate of institutional commitment and related variables over time on a sample of non-urgent undergraduates. Data for non-urgent clients at a University Counseling Center (UCC) were collected using on-line surveys over four time-periods. Within-time correlations generally showed that mental health concerns was negatively related to institutional commitment, while counseling help belief was positively related. Institutional commitment is defined as a student feeling that he or she selected the right institution to attend Using a smaller sample, i.e., n = 15, of complete-data clients matched-over-time, overall level of mental health concerns significantly declined, while institutional commitment significantly increased. Counseling help belief decreased from Time 1 to Time 2 but then increased over time. Scientifically demonstrating to higher-level University administration that counseling over time can positively influence undergraduates' institutional commitment can help the UCC to increase its allocation of university-based resources to keep pace with non-urgent client demands.

Keywords: counseling impact, institutional commitment, mental health concerns, non-urgent undergraduates, unlimited sessions

1. Introduction

The 2016 annual report by the Center for Collegiate Mental Health (CCMH, 2017), which collected data across 139 University Counseling Centers (UCCs), noted a continuing increase in the demand for mental health services. A supplemental survey of CCMH (2017) UCCs also found that, "on average UCCs are providing 28% more 'rapid-access' service hours per client and 7.6% fewer 'routine' service hours per client over the last six years" (p. 3). The report goes onto say that:

This shift, especially in centers with funding models that don't increase with demand, may impact the availability of routine services after the initial contact. Therefore, colleges and universities are encouraged to specifically attend to routine treatment capacity (not just the speed of initial response) as they consider responding to growing demand for mental health services in higher education (CCMH, 2017, p. 3).

The general purpose of this study was to examine counseling services as a correlate of institutional commitment and related variables over time on a sample of non-urgent undergraduates. This is an important issue to consider as UCCs compete for university resources with other campus stakeholders (Hunt, Watkins, & Eisenberg, 2012).

1.1 Why Focus on Institutional Commitment?

UCCs have typically reported research focused on evaluating their services, including: client satisfaction (Maffini & Toth, 2017), as well as other standardized mental health measures such as the: Behavioral Health Measure or BMH (Kopta et al., 2014), and Counseling Center Assessment of Psychological Symptoms or CCAPS (Brunner, Wallace, Reymann, Sellers, & McCabe, 2014). While gathering such mental health data is certainly important, 20 years ago Corazzini (1997, p. 380) noted that "the future of a counseling center is more promising when counseling is seen in relation to the academic mission". Archer and Cooper (1998) argued that providing services that were integrated with academic goals is one of the great challenges for a UCC. In addition to academic performance, fostering student commitment to the institution, via their engagement, is generally an important component of a

university/college's academic mission (Kezar & Kinzie, 2006). As an example of meeting this challenge, Coll and Stewart (2002) found that counseling services increased College of Education undergraduates' academic integration and social integration, as well as their teaching ability confidence. Research has long-argued (Tinto, 1975) that two forms of commitment—goal and institutional, are the immediate antecedents of undergraduate dropout decisions. Empirical research (Bean, 1980; Reason, 2009) has found that institutional commitment, i.e., feeling that you selected the right institution to attend (Davidson, Beck, & Milligan, 2009), is the most proximal antecedent of undergraduate retention. While research has shown that counseling has increased student retention (Lee, Olson, Locke, Testa Michelson, & Odes, 2009; Wilson, Mason, & Ewing, 1997), there has been less research investigating the relationship of counseling to client institutional commitment.

1.2 Correlates of Institutional Commitment

As noted above, institutional commitment would be expected to have a direct impact on retention rates. Given the difficulty of collecting student records, "intent to graduate" has often been used as a proxy for retention (e.g., Davidson et al., 2009; Hausmann, Ward Schofield, & Woods, 2007). When comparing Not-In-Counseling (NIC) with Initial Counseling Session (ICS), i.e., post-triage, pre-intake non-urgent undergraduate samples, Blau, DiMino, DeMaria, Beverly and Chessler (in press-a) found that institutional commitment was positively related to and explained additional significant variance in intent to graduate beyond controlled-for demographic variables (i.e., gender, ethnic background, residential status, parents' education, self-report GPA, and age) for both samples.

Why would counseling have a relationship to clients' institutional commitment? Using various pre-counseling and post-counseling measures, Brunner et al. (2014) called today's undergraduates the "most stressed" generation of college students (p. 264). When simultaneously comparing in-counseling versus not-in-counseling undergraduate samples, researchers have reported that students who are in-counseling have lower perceived well-being and mental health (Green, Lowry, & Kopta, 2003). College students with higher perceived psychological distress were less likely to persist towards graduation (Kitzrow, 2009). The perceived responsiveness of the institution, i.e., how/when it takes care of/attends to its' students, may be acutely significant when a student feels vulnerable and in need of help (Mowbray et al., 2006). For such a student, the UCC therapist can partially represent the institution, and the therapeutic alliance, i.e., collaborative and affective bond between therapist and patient, has been shown to affect subsequent treatment improvement (Bergin & Garfield, 2013). Pinkerton, Tally, and Cooper (2009) noted the benefits of open-ended psychotherapy, which allowed the client to work through developmental milestones over a period of time (e.g., leaving home, selecting a major, defining a career), as well as upheavals (e.g., break up of an intimate relationship, family death). This research collectively suggests that level of mental distress, mental health concerns, belief in counseling helping, and intent to graduate are correlates of institutional commitment.

1.3 Generally Brief Nature of College Counseling

Given the increased demand for mental health services on many university campuses (Gallagher, 2014), and the typical way that college students use such services, counseling is typically of short duration. Using the data gathered from 1,698 college student clients across 42 universities, Draper, Jennings, Baron, Ozgur, and Shankar (2002) found that the average number of sessions per client was 3.3. In another review, Ghetie (2007) noted a median of 4-5 counseling sessions and a mode between 1 and 2 sessions. Mahon et al. (2015) found that 37/124 (30%) of undergraduate clients completed a minimum of three counseling sessions, with the remaining 87 either never returning for a second session or dropping out after two sessions. The 2016 CCMH report (CCMH, 2017) noted a mode of 1 for the number of appointments and an average of 5.84 sessions, while the 2016 Association for University and College Counseling Directors or AUCCD report (AUCCD, 2016) reported an average of 5.53 sessions/client.

While it can be very challenging to build a longer-term sample of undergraduates receiving counseling services, prior research suggests that even a short-term intervention can increase important undergraduate outcomes. Mahon et al. (2015) found a significant decrease in counselor-reported client symptom severity. Blau et al. (2016) found that the means of social connectedness and life satisfaction significantly increased over time for 28 non-urgent undergraduates, with a median of four counseling sessions (after initial triage assessment).

1.4 Longer-Duration College Counseling

Surette and Shier (2017) worked with 102 clients for a mean of seven sessions over an average of 18 weeks, and assessed the outcomes of somatization, depression, and anxiety. They found that duration of therapy, positive life events, and intimate relationship supports positively influenced outcomes, while both an increased number of intervention approaches and sessions each had a negative influence. In a two-part study over three years, Lucas

(2012) first compared three groups of non-urgent undergraduate clients: no-show first session (n = 492); stopped before improvement (n = 396), and counseling completers (n = 1,042), and found that completers compared to the other two groups, reported at intake higher levels of confidence that counseling will help, degree of well-being, and life functioning. Study 2 (Lucas, 2012) focused on students (n = 75) who received at least eight counseling sessions and found that students who rated themselves as highly distressed (n = 47) versus less distressed (n = 28) at intake reported the most benefits. Kopta et al. (2014) used a data set of 13,803 clients on the Behavioral Health Measure-20 (BMH-20) across 23 UCCs in the United States from 2006-11 to show that client treatment benefit on the BMH-20 peaked between 7 to 10 sessions.

Using a smaller sample of 20 matched-over-three time-periods non-urgent clients, Blau, DiMino, Abreu, and LeLeux-LaBarge (in press-b) found that levels of mental distress and mental health concerns significantly declined, while life satisfaction significantly increased. The three time periods were over a period of ten weeks (range 8 to 12 weeks), with a median of six counseling sessions after intake (range less than 5 to 12). The current study focused on a different outcome than (Blau et al., in press-b), i.e., institutional commitment, and it is also different from Blau et al. (in press-a) by testing a new independent variable (counseling help belief). The current study looked at the relationship of counseling with a greater number of sessions over a longer time period then these two prior (Blau et al., in press-a; Blau et al., in press-b) studies. Prior research with UCC clients (Wolgast, Lambert, & Puschner, 2004) has shown that the number of sessions affected clinically significant change, with 14 sessions required for 51% of clients to meet this change criteria. Clinically significant change was defined by "when a client's final Outcome Questionnaire or OQ-45 (Lambert et al., 1996) was equal or less than 63 and had decreased by at least 14 points from the intake session" (Wolgast et al., p. 21). The OQ-45 measures symptoms of psychopathology, interpersonal problems, and social role functioning difficulties. With a focus on institutional commitment as the main study variable, the four hypotheses for this study were:

H1—Level of mental distress and mental health concerns will be negatively related, while counseling help belief will be positively related, to institutional commitment

H2—Institutional commitment is positively related to intent to graduate

H3—Brief counseling will increase intent to graduate

H4—Over a longer period of time counseling will reduce the level of mental distress and mental health concerns and increase counseling help belief and institutional commitment

2. Method

2.1 Samples and Procedures

Four undergraduate samples based on length of time in counseling were collected in this study. All data were voluntarily collected using on-line surveys at the main campus of a large state-supported urban university's UCC in the Mid-Atlantic region of the United States. The first sample is labeled as *initial counseling session sample or Time 1*. All three subsequent samples were derived from this initial sample and are labeled *Time 2* and *Time 3* and *Time 4*. However, for any given client, the nature of their mental issue, as well as client/therapist availability did not allow for a standardized number of sessions or time periods to represent when clients filled out the subsequent surveys. Given this non-standardized time-period and number of counseling sessions provided for participants before taking subsequent surveys, a range of sessions and time-periods were used (reported below).

Data were collected for 332 undergraduates who registered for counseling services at the UCC between the Fall of 2013 through the Spring of 2017. Institutional Review Board approval was given for all data collection, and data went directly into Qualtrics' data bases. These undergraduate students had been designated at their initial counseling session (i.e., an in-person triage interview), as non-urgent (non-suicidal or no risk to self/others) by the triage therapist. Students were asked by the triage therapist if they wished to voluntarily participate in a research study looking at the impact of counseling on student outcomes. The *Time 1* clients took their survey after triage but before beginning counseling (i.e., intake). The time span between triage and intake could be anywhere from that week to over two to three weeks, depending on current client demand for services and therapist availability. Sample demographics for the Time 1 complete-data sample are reported below.

Undergraduate clients who filled out the initial survey (*Time 1*) were given the option of listing the last four digits of their nine-digit university identity number so that their responses, while receiving counseling, could be tracked over time while their individual identities would remain protected. Clients could take the Time 1 survey either at the UCC's self-help center or at home using a survey link. Using their email addresses recorded at triage, clients were contacted after their triage appointment at four week intervals, i.e., at Time 2 (four weeks from triage), Time

3 (eight weeks from triage), and Time 4 (12 weeks from triage). A general email was sent out to all post triage clients on a rolling weekly basis during a semester. As a check, clients were asked what survey they were filling out (Time 1, Time 2, Time 3, Time 4). Given the volume of clients involved, and the anonymity of respondents (using only the last 4-digits), it was not determined if a particular client had stopped coming to the UCC.

Of the 332 initial participants, 211 (64%) filled out all the Time 1 survey data and are designated as *Time 1*. A sample of n = 64/211 (39%) filled out a second survey at *Time 2*. This *Time 2* sample reported a median of three counseling sessions after intake (range 1 to 8) over a median period of six weeks (range 4 to 10 weeks). Of the *Time 1* clients, N = 48/211 (23%) filled out the third survey or *Time 3*. This *Time 3* sample reported a median of six counseling sessions after intake (range less than 5 to 12) over a median period of ten weeks (range 8 to 12 weeks). Finally, for *Time 4*, 31/211 (15%) filled out the fourth survey, reporting a median of nine counseling sessions after intake (range less than 9 to 15) over a median period of fourteen weeks (range 12 to 20 weeks). Thus there is a client participation rate decline from N = 211 (*Time 1*) to N = 64 (*Time* 2) to N = 48 (*Time 3*) to N = 31 (*Time 4*). In order to improve the longitudinal response rate a $50 random-drawing lottery for matched ID repeat-respondents was approved by the IRB part-way through the data collection process. However, this incentive had little impact in improving the matched respondent participation rate. Across all four time periods there were only 15 clients who matched across time, when using their four-digit identification number. This four-time match sample (N = 15) will be called *the complete sample*.

2.2 Measures

Respondent demographics. Demographics were collected only at Time 1. Data collected were: main mental health issue for presenting to counseling, gender (male/female/transgender), ethnic background (non-white/white), status entering university (no transfer/transfer, i.e., did the student start at the university as a freshman versus transfer in at some later point), current residential status (non-commuter/commuter), parents' highest education level, self-reported Grade Point Average (GPA), age, referral source for counseling (self/other), days waited for intake (after triage) and status as a full-time student (taking at least 12 credits/semester).

Level of mental distress, mental health concerns, counseling help belief, institutional commitment, and intent to graduate. Unless otherwise noted, all items were measured using a seven-point Likert scale, where 1 = strongly disagree to 7 = strongly agree. These variables were measured at all four times, except for intent to graduate which was only measured at Time 1 and 2. *Level of mental distress* was measured by asking clients to "rate your current level of mental distress that brought you to counseling". An eight-point Likert response scale was used, where: 1 = none, 2 = very little, 3 = a little, 4 = moderate, 5 = a lot, 6 = severe, 7 = very severe, 8 = unbearable. *Mental health concerns* were measured using six items. Items were "I have recently: (1) lost a lot of sleep over worrying, (2) felt constantly under some type of strain, (3) felt that I could not overcome my difficulties, (4) been feeling unhappy or depressed, (5) been losing confidence in myself and (6) been thinking of myself as a worthless person." These six items represent the Symptoms of Mental Disorder factor in the short-form of the General Health Questionnaire (Hu, Stewart-Brown, Twigg, & Weich, 2007). The 4-point response scale used by Hu et al. (2007, p. 1006) was expanded to a 7-point response scale to keep response scale consistency. The higher the total score the greater one's mental health concerns. Blau et al. (in press-a) found the Cronbach's alpha was .91 NIC sample and .80 for the ICS sample.

Belief in counseling helping (*Counseling Help Belief*) was measured using six study-specific items, "I believe that counseling will help me to: (1) better cope with difficult situations I have, (2) more successfully deal with the life challenges I face, (3) interact with others more successfully, (4) learn how to better manage my negative thinking, (5) understand how my prior learned thought patterns may be contributing to my current concerns, and (6) reaffirm my inner personal strengths." These items were partially based on the qualitative analyses about what clients felt was helpful about counseling, e.g., client resolutions, new perspectives, gaining knowledge, emotional relief (Paulson, Truscott, & Stuart, 1999). Previous research (Blau et al., 2015) found an internal consistency reliability of this scale (coefficient alpha) was .92. *Institutional commitment* was measured using three of four items (Davidson et al., 2009). Items were: "(1) I am sure University X is the right place for me, (2) I am confident that I made the right decision to attend University X, and (3) I have no plans to transfer/leave University X." The fourth item was not used because it mentions "earning a degree", which potentially overlaps with intent to graduate. The Cronbach's alpha was .86 for both the NIC and ICS samples (Blau et al., in press-a). *Intent to graduate* was measured using three items. Items were: "(1) I intend to persist in getting a college degree, (2) I am confident that I will graduate with a degree, and (3) no matter what it takes, I am committed to getting my college degree." These items were adapted from prior research (Davidson et al., 2009; Hausmann et al.,

2007). The higher the total score the greater one's intent to graduate. The Cronbach's alpha was .92 for the NIC sample and .88 for the ICS sample (Blau et al., in press-a).

2.3 Data Analyses

All data analyses were done using SPSS-PC (SPSS, 2013). The two hypotheses (H1 & H2) were tested using within-time period Pearson correlation analyses for complete data samples on the variables involved. The third hypothesis (H3), involving only intent to graduate over two time periods, was tested using a paired samples t-test. The fourth hypothesis (H4) was tested using the General Linear Model (GLM) for each of the four variables across four time periods. The GLM was appropriate for a repeated measures design (Stevens, 1996). There was no homogeneity of variance test with the GLM because only one group, i.e., the *complete sample*, was used. However, Mauchly's sphericity test, which is appropriate for three or more levels of a repeated measure factor (i.e., time), was calculated prior to testing between group differences (Stevens, 1996). A non-significant test indicated that the variances of the differences between all possible pairs of within-subject conditions (i.e., levels of independent variable or four different times) was equal and the subsequent multivariate F test was not inflated. For each GLM test, Mauchly's sphericity test was not significant, supporting the subsequent multivariate F test reported. If a significant multivariate F was found, the least squares difference post hoc test between groups was then used (Stevens, 1996). Since direction was specified in H4, an *a priori* one-tail test is justified (Stevens, 1996). Effect size for the largest significant difference between groups within a variable was also reported (Cohen, 1988).

3. Results

3.1 Demographic Comparison of Complete to Initial Sample

Table 1 shows a demographic comparison of the complete-data *Time 1* sample (n = 211) to the *Time 4* sample (n = 31). Overall, there was general consistency between the two samples, although the percentage of males by Time 4 had noticeably dropped while, the percentages of female and transgender clients had increased. Sample participants reporting anxiety and depression as the two main mental health issues for coming to the UCC was consistent with the AUCCD (2016) and CCMH (2017) reports. However, the complete data respondent percentage reporting anxiety at Time 4 had risen considerably from Time 1. Anxiety seemed to be a stronger mental health issue for non-urgent clients staying longer in counseling. The UCC sample participant base of primarily white female was also consistent with these two reports.

Table 1. Demographic variable comparison for Time 1 versus Time 4 samples

Variable	Time 1 (N = 211)	Time 4 (N = 31)
Main Mental Health Issue		
Anxiety	39%	59%
Depression	33%	27%
Gender		
Male	30%	5%
Female	68%	86%
Transgender	2%	9%
Ethnic background		
Non-white	26%	32%
White	74%	68%
Status		
No transfer	63%	59%
Transfer	37%	41%
Residential status		
Non-commuter	69%	68%
Commuter	31%	32%

Parents' education		
Less than 4-year college degree	30%	45%
At least 4-year college degree	70%	55%
Self-reported GPA (range)	3.4 (2.0-4.0)	3.5 (2.9-4.0)
Age, M (range)	21 (18-29)	22 (18-30)
Referral Source		
Self	61%	59%
Other (e.g., family, friend)	39%	41%
Mean Days Waited for Intake	15	14
(range)	(1-41)	(1-30)
Full-time Student (12 plus credits)	97%	96%

3.2 Testing the Hypotheses

To test both H1 and H2 within time correlational analyses are shown in Table 2. For H1, results show that level of mental distress was not related to Institutional Commitment (IC), while mental health concerns was significantly negatively related to IC for two of four time periods. Counseling help belief was positively related to IC for three of four time periods. Collectively these results provide partial support for H1. For H2, IC was positively related to intent to graduate within both Time 1 and 2. In addition, the correlation of Time 1 institutional commitment to Time 2 intent to graduate was significant, i.e., $\underline{r} = .30$, $n = 64$, $p < .01$ (one-tailed), which supports H2.

Table 2. Means, standard deviations, and correlations among level of mental distress, mental health concerns, counseling help belief and institutional within four separate times[a]

Time 1 (N = 211)							
Variable	M	SD	1	2	3	4	5
1. Level of Mental Distress	4.85	1.09	----				
2. Mental Health Concerns	5.15	1.08	.49**	-----			
3. Counseling Help Belief	6.02	.74	-.07	-.01	-----		
4. Institutional Commitment	5.54	1.41	-.12	-.06	.17*	-----	
5. Intent to Graduate[a]	6.42	.83	-.15*	-.24**	.27**	.35**	----
Time 2 (N = 64)							
Variable	M	SD	1	2	3	4	5
1. Level of Mental Distress	4.42	1.33	-----				
2. Mental Health Concerns	4.79	1.35	.54**	-----			
3. Counseling Help Belief	5.12	1.15	-.22	-.42**	-----		
4. Institutional Commitment	5.63	1.42	-.07	-.31*	.28*	-----	
5. Intent to Graduate[a]	6.67	.50	.14	-.10	.32*	.24*	-----
Time 3 (N = 48)							
Variable	M	SD	1	2	3	4	
1. Level of Mental Distress	3.96	1.35	-----				
2. Mental Health Concerns	4.15	1.29	.51**				
3. Counseling Help Belief	5.50	1.19	-.29*	-.35*			
4. Institutional Commitment	5.42	1.66	-.03	-.25	.03	-----	

Time 4 (N = 31)						
Variable	M	SD	1	2	3	4
1. Level of Mental Distress	3.65	1.23	-----			
2. Mental Health Concerns	4.23	1.27	.47**			
3. Counseling Help Belief	5.60	1.22	-.42*	-.37*		
4. Institutional Commitment	5.83	1.31	-.19	-.38*	.35*	-----

Note. * p < .05; ** p < .01 (two-tailed). Listwise deletion of data within each time.

[a] Intent to Graduate data only collected at Time 1 and Time 2.

The paired sample t-test for H3, did not show a significant change in intent to graduate means from time 1 (M = 6.63) to time 2 (M = 6.70), t(65) = -1.26, p = .11. Thus, H3 is not supported. For H4, the results are shown in Table 3. Significant overall changes (multivariate F) in the *complete sample* were found for three of the four variables tested: mental health concerns, counseling help belief, and institutional commitment. Mental health concerns declined, and the mean difference decline between Time 1 (M = 5.52) and Time 3 (M = 4.30) was the largest. For the largest difference in groups, the effect size was 1.15 (Cohen, 1988), which is classified as large. Counseling help belief started out strong (M = 6.20) but then dropped at Time 2 (4.82) and then rebounded at Time 3 (M = 5.54) and Time 4 (M = 5.64). The effect size of .75 (large) was calculated for improvement from Time 2 to Time 4. For institutional commitment, there was drop from Time 1 to Time 2 but then a steady increase at Time 3 and Time 4. The increase from Time 2 (M = 4.69) to Time 4 (5.24) was significant, and the effect size was .44 (medium). Although the multivariate F for level of mental distress was not significant (F = 1.79), the descriptive data shows cell mean decreases over time. Overall, there was partial support for H4.

Table 3. General linear model testing significance of outcome variable, post hoc differences within subgroups, and largest effect size for complete sample

Outcome Variable	Level of Mental Distress				Mental Health Concerns				Counseling Help Belief				Institutional Commitment			
Multivariate F	F = 1.79				F = 3.38*				F = 4.85*				F = 3.85*			
Cell Means	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
	4.93	4.57	3.64	3.63	5.52[a]	5.28[a]	4.30[b]	4.35[b]	6.20[a]	4.82[b]	5.54[a]	5.64[a]	4.79	4.69[a]	5.00	5.24[b]
Effect Size[d]					*d* = 1.15				*d* = .75				*d* = .44			

Note. N = 15; T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4.

Response scale for all outcome variables, 1 = strongly disagree to 7 = strongly agree.

[a,b]Within each variable, cell means that do not share the same superscript are significantly different at the 0.05 level by the Least Significant Difference post hoc test.

[d]Cohen's (1988) effect size, where $d = M_1\text{-}M_2/\sqrt{}$pooled SD for largest significant difference within outcome variable groups, except for counseling Help Belief, where *d* is based on the difference between T2 and T4.

* p < .05 (one-tailed).

4. Discussion

To the authors' knowledge, the research design of this study is rare in finding significant correlations among study variables across four time periods for a non-urgent sample of students receiving counseling services; as well as significant changes, with large effect sizes (Cohen, 1988), in two of four variables for the complete counseling sample (mental health concerns, counseling help belief). The decrease in mental health concerns is consistent with clinician rated change in client distress reported in the AUCCD (2016). The robustness of these two findings, despite the loss in client participation over time, is encouraging. The drop in counseling help belief from Time 1 to Time 2 may have been partially due to clients realizing that their mental health concerns were not as "episodic" as

typically found for college students (Webb & Widseth, 1988). Despite this drop in counseling help belief, these complete-sample clients continued in therapy, and their belief increased over time.

These overall improvement results are also consistent with Lucas (2012) who found significant improvement in the "highly distressed" undergraduate client sub-sample between intake and their eighth session, and Wolgast et al. (2004) who showed that 14 sessions were needed for clients to show clinically significant change. Such research supports having the resources available in a UCC to avoid "time limited" treatment (Ghetie, 2007). Many college students successfully engage in short-term treatment (Mahon et al., 2015). However, knowing that additional sessions can lead to further benefits in terms of reduced mental health concerns and increased counseling help belief should help to justify a UCC having the resources in place, such as adequate staffing, to avoid having a fixed number or cap of sessions with non-urgent college students (AUCCD, 2016; CCMH, 2017; Lucas, 2012).

What is more unique about this study is the finding that counseling had an overall positive relationship to institutional commitment for the complete sample, and the more limited support that institutional commitment had a positive correlation to intent to graduate. Institutional commitment has been defined as students' feeling that they made the right institutional choice (Davidson et al., 2009), and more general college student research has shown that it is the most proximal antecedent of undergraduate retention (Bean, 1980; Reason, 2009). Focusing on institutional commitment is consistent with prior calls to have UCCs studying variables consistent with more general academic missions and goals (Archer & Cooper, 1998; Corazzini, 1997). Prior research (Kezar & Kinsey, 2006) has noted that fostering student commitment to the institution, via their engagement, is generally an important component of a university/college's academic mission. UCCs finding a way to not only show that counseling has mental health benefits, but also benefits for the university/college should allow the UCC to have a more competitive advantage as it competes with other university stakeholders for resources (Hunt et al., 2012).

Two scales, mental health concerns and institutional commitment, were successfully used as shortened versions of longer scales. In addition, a new counseling help belief scale was utilized. All three multi-item scales demonstrated good scale reliabilities. Prior research working with college student samples has suggested that shorter survey length can help with a higher response rate (Fan & Yan, 2010).

4.1 Study Limitations

One study limitation was that intent to graduate data were only collected for the first two time periods. In addition, only non-urgent clients were sampled, and all data were self-reported. However, the biggest limitations of the current study were the generally poor initial response and then the loss in client sample size over time. There were only 15 participants in the complete-data sample. Although such client loss is consistent with prior research (e.g., Draper et al., 2002; Ghetie, 2007; Lucas, 2012; Mahon et al., 2015), it was none-the-less discouraging. The four-year data collection has stopped due to several reasons, including the UCC moving to a new location and some key staff turnover, but general lack of client response was the primary issue. There were a number of reasons illustrating the difficulty of a single UCC collecting a large longitudinal data set. Clearly, many students when given a choice to participate in the research study by the triage therapist said "no". Despite assurances of confidentiality, clients may not have wanted their answers tracked longitudinally. In addition, for those who started the study, archival data from the UCC study site showed that, consistent with the CCMH (2017) report, one is the most frequent number of sessions for a client. Further archival data from the UCC site showed that by 4 sessions (triage plus three sessions after) over 50% of the non-urgent clients had stopped counseling. Despite the general demographic consistency between the Time 1 and Time 4 samples, the loss in respondents over time is an internal validity threat, complicated by the variability in number of counseling sessions (and time ranges) between the fixed data collection time intervals after triage. In addition, the percentage of clients reporting anxiety as their main mental health issue increased considerably from Time 1 to Time 4, and the percentage of female (male) clients increased (decreased).

University growth in student enrollment over the last four years has led to more on-campus and close-to-campus housing making the UCC's counseling services more accessible. This student-growth and greater accessibility has led to increased demand for the UCC's counseling services over the past several years (e.g., 34% increase in intakes, 37% increase in individual counseling), without the needed increase in UCC professional staff to meet the recommended ICAS ratio of 1 FTE professional staff to every 1,000 to 1,500 students (IACS, 2010). For such an under-staffed UCC, voluntary client research would be expected to have a lower priority. Despite the UCC Director and several staff being supportive, it is probably safe to assume that many already-stressed triage therapists never asked students initially if they were interested in study participation. As Cooper and Archer (2002) noted 15 years ago, "the mission of counseling centers is primarily service, and the reward structure for individuals

and centers does not typically encourage research and scholarly writing" (p. 57). In addition, the lead author was not part of the UCC professional staff, but instead an "outside" faculty member. Although in-person progress reports on data collection during each regular semester were given by the lead author at UCC staff meetings, as well as other periodic updates, research project visibility to the staff undoubtedly suffered due to higher priorities. Offering a money-incentive, as noted in the Methods section, part-way through the data collection did not improve client participation. One solution to the difficulty of a single UCC collecting sufficient data is for the UCC to become part of a larger research network so that common data collection from multiple UCCs can be aggregated (Kopta et al., 2014).

4.2 Implications for Practicing College Clinicians and Resourcing UCCs

For practicing UCC counselors, the findings that over time counseling help belief rebounded, as well as institutional commitment increased, both support the therapist not having to cap the number of sessions with a non-urgent client (AUCCD, 2016). These findings reinforce the benefits of an open-ended psychotherapy process (Pinkerton et al., 2009). As part of this process, the study results also support the therapist being able to spread out counseling sessions over time if necessary (Surette & Shier, 2017).

Scientifically demonstrating to University administration that the on-campus UCC site can help undergraduates in distress should ideally help the UCC to increase its allocation of university-based resources (Bishop, 2010; Hunt et al., 2012). Such research efforts need to be continued as UCCs compete with other university stakeholders for resources (Castonguay, Locke, & Hayes, 2011). However, beyond more traditional measures such as the BMH (Kopta et al., 2014) or CCAPS (Brunner et al., 2014), UCC-site research efforts which also incorporate variables more directly linked to the academic mission of a university or college, are needed. It may be necessary for a particular UCC to demonstrate to its University Administration that counseling benefits the University's student retention and immediate retention-antecedents, such as institutional commitment. By showing that a UCC not only reduces its' students mental health concerns, but also helps to meet a university's academic goals, this should help the UCC to get the resources needed to keep up with increasing routine as well as non-routine student mental health demands.

References

AUCCD. (2016). *The association for university and college counseling center directors annual Survey*. Retrieved from http://www.aucccd.org/assets/.../aucccd%202016%20monograph%20-%20public.pdf

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, *12*(2), 155-187. https://doi.org/10.1007/BF00976194

Bergin, A., & Garfield, S. (2013). *Handbook of psychotherapy and behavior change* (6th ed.). New York, N.Y.: John Wiley & Sons.

Bishop, J. B. (2010). The counseling center: An undervalued resource in recruitment, retention and risk management. *Journal of College Student Psychotherapy*, *24*, 248-260. https://doi.org/10.1080/87568225.2010.509219

Blau, G., DiMino, J., Abreu, I., & LeLeux-LaBarge, K. (in press-b). Testing the impact of counseling over time on non-urgent undergraduate life satisfaction. *Journal of Educational and Developmental Psychology*, *7*(2), 59. https://doi.org/10.5539/jedp.v7n2p59

Blau, G., DiMino, J., DeMaria Jr., P., Beverly, C., & Chessler, M. (in press-a). Mental health concerns' impact on graduation intent and improvement for brief counseling. *Journal of Assessment and Institutional Effectiveness*.

Blau, G., DiMino, J., DeMaria, P. A. Jr., Beverly, C., Chessler, M., & Drennan, R. Jr. (2016). Social connectedness and life satisfaction: Comparing mean levels for two undergraduate samples and testing for improvement based on brief counseling. *Journal of American College Health*, *64*(8), 585-592. https://doi.org/10.1080/07448481.2016.1207645

Blau, G., DiMino, J., Sheridan, N., Stein, A., Casper, S., Chessler, M. & Beverly, C. (2015). Wait time for counseling affecting perceived stigma and attitude toward the University. *College Student Journal*, *49*(2), 280-290.

Brunner, J. L., Wallace, D. L., Reymann, L. S., Sellers, J.-J., & McCabe, A. G. (2014). College counseling today: Contemporary students and how counseling centers meet their needs. *Journal of College Student Psychotherapy*, *28*, 257-324. https://doi.org/10.1080/87568225.2014.948770

Castonguay, L. G., Locke, B. D., & Hayes, J. A. (2011). The Center for Collegiate Mental Health: An example of a practice-research network in university counseling centers. *Journal of College Student Psychotherapy*, *25*(2), 105-119. https://doi.org/10.1080/87568225.2011.556929

CCMH. (2017). *Center for collegiate mental health: 2016 Annual report*. Penn State (Publication No. STA 17-74). Retrieved from https://sites.psu.edu/ccmh/files/2017/01/2016-Annual-Report-FINAL_2016_01_09-1gc2hj6.pdf

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Coll, K. M., & Stewart, R. A. (2002). Collaboration between counseling services and an academic program: An exploratory study of student outcome. *Journal of College Counseling*, *5*, 135-141. https://doi.org/10.1002/j.2161-1882.2002.tb00215.x

Cooper, S. E., & Archer, J. A. Jr. (2002). Evaluation and research in college counseling center contexts. *Journal of College Counseling*, *5*, 50-59. https://doi.org/10.1002/j.2161-1882.2002.tb00206.x

Corazzini, J. (1997). Using research to determine the efficacy and modes of treatment in University Counseling Centers: Comment on the Jobes, Jacoby, Cimbolic, and Hustead (1997) and Hayes (1997). *Journal of Counseling Psychology*, *44*(4), 378-380. https://doi.org/10.1037/0022-0167.44.4.378

Davidson, W. B., Beck, H. P., & Milligan, M. (2009). The College Persistence Questionnaire: Development and validation of an instrument that predicts student attrition. *Journal of College Student Development, 50*(4), 373-390. https://doi.org/10.1353/csd.0.0079

Draper, M. R., Jennings, J., Baron, A., Erdur, O., & Shankar, L. (2002). Time-limited counseling outcome in a nationwide college counseling center sample. *Journal of College Counseling*, *5*(1), 26-38. https://doi.org/10.1002/j.2161-1882.2002.tb00204.x

Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior*, *26*(2), 132-139. https://doi.org/10.1016/j.chb.2009.10.015

Gallagher, R. P. (2014). *National Survey of Counseling Center Directors*. Retrieved December 28, 2016, from http://www.iacsinc.org/

Ghetie, D. (2007). The debate over time-limited treatment in college counseling centers. *Journal of College Student Psychotherapy*, *22*(1), 41-61. https://doi.org/10.1300/J035v22n01_04

Green, J. L., Lowry, J. L., & Kopta, M. (2003). College students versus college counseling center clients. *Journal of College Student Psychotherapy*, *17*(4), 25-37. https://doi.org/10.1300/J035v17n04_05

Hausmann, L. R., Ward Schofield, J., & Woods, R. L. (2007). Sense of belonging as a predictor of intentions to persist among African American and White first-year college students. *Research in Higher Education*, *48*(7), 803-839. https://doi.org/10.1007/s11162-007-9052-9

IACS. (2010). *Accreditation for the university and college counseling services*. Alexandria, VA: International Association of Counseling Services, Inc.

Hu, Y. J., Stewart-Brown, S., Twigg, L., & Weich, S. (1997). Can the 12-item General Health Questionnaire be used to measure positive mental health? *Psychological Medicine*, *37*(7), 1005-1013. https://doi.org/10.1017/S0033291707009993

Hunt, J. B., Watkins, D., & Eisenberg, D. (2012). How do college campuses make decisions about allocating resources for student mental health? Findings from key participant interviews. *Journal of College Student Development*, *53*(6), 850-856. https://doi.org/10.1353/csd.2012.0083

Kezar, A., & Kinzie, J. (2006). Examining the ways institutions create student engagement: The role of mission. *Journal of College Student Development*, *47*(2), 149-172. https://doi.org/10.1353/csd.2006.0018

Kitzrow, M. A. (2009). The mental health needs of today's college students: Challenges and recommendations. *NASPA Journal*, *46*(4), 646-660. https://doi.org/10.2202/1949-6605.5037

Kopta, S. M., Petrik, M. L., Saunders, S. M., Mond, M., Hirsch, G., Kadison, R., & Raymond, D. (2014). The utility of an efficient outcomes assessment system at university counseling centers. *Journal of College Student Psychotherapy*, *28*, 97-116. https://doi.org/10.1080/87568225.2014.883876

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., & Reisinger, C. W. (1996). *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire).* American Professional Credentialing Services, LLC.

Lee, D., Olson, E. A., Locke, B., Testa Michelson, S., & Odes, E. (2009). The effects of college counseling services on academic performance and retention. *Journal of College Student Development*, *50*(3), 305-319. https://doi.org/10.1353/csd.0.0071

Lucas, M. S. (2012). Counseling on campus: Client persistence and progress. *Journal of College Student Psychotherapy*, *26*, 227-240. https://doi.org/10.1080/87568225.2012.685856

Mahon, M., Laux, J. M., McGuire Wise, S., Ritchie, M. H., Piazza, N. J., & Tiamiyu, M. (2015). Brief therapy at a university counseling center: Working alliance, readiness to change, and symptom severity. *Journal of College Counseling*, *18*, 223-243. https://doi.org/10.1002/jocc.12017

Maffini, C. S., & Toth, P. L. (2017). Quality improvement in university counseling centers. *Journal of College Student Psychotherapy*, *31*(1), 43-58. https://doi.org/10.1080/87568225.2016.1254000

Mowbray, C., Megivern, D., Mandiberg, J., Strauss, S., Stein, C., Collins, K. … Lett, R. (2006). Campus mental health services: Recommendations for change. *American Journal of Orthopsychiatry*, *76*(2), 226-237. https://doi.org/10.1037/0002-9432.76.2.226

Paulson, B., Truscott, D., & Stuart, J. (1999). Clients' perceptions of helpful experiences in counseling. *Journal of Counseling Psychology*, *46*(3), 317-324. https://doi.org/10.1037/0022-0167.46.3.317

Pinkerton, R., Talley, J. E., & Cooper, S. L. (2009). Reflections on individual psychotherapy with university students: What seems to work. *Journal of College Student Psychotherapy*, *23*, 153-171. https://doi.org/10.1080/87568220902793830

Reason, R. D. (2009). Student variables that predict retention: Recent research and new developments. *NASPA Journal*, *46*(3), 482-501. https://doi.org/10.2202/1949-6605.5022

SPSS. (2013). *Statistical package for the social sciences* (Version 22). Armonk, NY: IBM.

Stevens, J. (1996). *Applied Multivariate Statistics for the Social Sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Surette, T. E., & Shier, M. L. (2017). A common factors approach to supporting university students experiencing psychological distress. *Journal of College Student Psychotherapy*, *31*(2), 112-131. https://doi.org/10.1080/87568225.2016.1248233

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*(1), 89-125. https://doi.org/10.3102/00346543045001089

Webb, R., & Widseth, J. (1988). Facilitating students' going into and stepping back from their inner worlds: Psychotherapy and the college student. *Journal of College Student Psychotherapy*, *3*(1), 5-15. https://doi.org/10.1300/J035v03n01_02

Wilson, S. B., Mason, T. W., & Ewing, M. J. (1997). Evaluating the impact of receiving university-based counseling services on student retention. *Journal of Counseling Psychology*, *44*(3), 316-320. https://doi.org/10.1037/0022-0167.44.3.316

Wolgast, B. M., Lambert, M. J., & Puschner, B. (2004). The dose-response relationship at a college counseling center. *Journal of College Student Psychotherapy*, *18*(2), 15-29. https://doi.org/10.1300/J035v18n02_03

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Construct Validation of the Motivated Strategies for Learning Questionnaire in a Singapore High School Sample

Chiu Wai Chow[1,2] & Elaine Chapman[1]

[1] Graduate School of Education, The University of Western Australia, Perth, Australia

[2] Hwa Chong Institution, Singapore

Correspondence: Chiu Wai Chow, Education Technology Department, Hwa Chong Institution, 661 Bukit Timah Road, 269734, Singapore. E-mail: chowcw@hci.edu.sg

Received: August 9, 201 Accepted: August 22, 2017 Online Published: September 20, 2017

doi:10.5539/jedp.v7n2p107 URL: http://doi.org/10.5539/jedp.v7n2p107

Abstract

In this study, the construct validity of the Motivated Strategies for Learning Questionnaire (MSLQ) was assessed. Participants were 441 Year 11 students in Singapore. Three separate confirmatory factor analyses were conducted for each section of the MSLQ (motivation and learning strategies). Results indicated that the original factor structures proposed by the instrument developers produced the best model fit. Cronbach α coefficients were also acceptable for all but one of the individual scales. Correlations with the Revised Learning Process Questionnaire—Two Factor and physics achievement scores also aligned with the theoretical basis of the MSLQ. These results confirmed the potential utility of this instrument for assessing the motivation and learning strategies of secondary students in Singapore.

Keywords: MSLQ, validation, motivation, learning strategies

1. Introduction

The 21st Century Competency Framework developed by the Singapore Ministry of Education in 2010 places central importance on the development of students' self-regulatory skills (Ministry of Education Singapore, 2010). The term self-regulation, when applied to learning, refers to the proactive process whereby learners set goals for their learning, actively monitor their progress, and regulate their cognition, motivation and behavior in order to achieve their learning goals (Pintrich, 2000). Research has indicated that individuals with higher levels of self-regulation are not only more successful in schools, but also in other aspects of their lives. In addition to performing better academically, these students tend also to achieve greater success in their careers, and enjoy better health, than those with lower levels of self-regulation (Bandura, 1982; Baumeister, Heatherton, & Tice, 1994; Boekaertz, Pintrich, & Zeidner, 2005; Locke & Latham, 2002; Mischel, Shoda, & Rodriguez, 1989).

The Ministry of Education's 2010 initiative points to the need for a validated instrument that can be used to assess the self-regulation levels of secondary level students in Singapore. One of the most widely used instruments for assessing students' self-regulated learning is the Motivated Strategies for Learning Questionnaire (MSLQ: Pintrich, Smith, García, & McKeachie, 1991, 1993). Duncan and McKeachie (2005) identified 55 empirical studies that had employed either the entire MSLQ or part of it within just a five-year period (2000-2004). At the college level, the instrument has been applied in studies across Western countries (e.g., Campbell, 2001; McKenzie & Gow, 2004; Suárez, González, & Valle, 2001) and other contexts (e.g., Cheung, Rudowicz, Lang, Yue, & Kwan, 2001; Ostovar & Khayyer, 2004). In 2011, a meta-analysis by Credé and Phillips (2011) identified 67 studies that had used the MSLQ across 19,900 college students. The MSLQ has also been used at the secondary level in different countries, including America (e.g., Liu, 2003), Germany (e.g., Neber & Heller, 2002), Hong Kong (e.g., Sachs, Law, & Chan, 2002), Israel (e.g., Eshel & Kohavi, 2003), Korea (e.g., Bong, 2001) and Turkey (e.g., Andreou, 2004).

The MSLQ is divided into two distinct sections. The motivation section assesses three main constructs (García & Pintrich, 1995): values (i.e., students' perceptions of the importance and interest of tasks), expectancy beliefs (i.e., students' beliefs about their task competency), and affect (i.e., students' emotional reactions to learning tasks). Six scales are used to assess these three constructs: Intrinsic Goal Orientation, Extrinsic Goal Orientation,

Task Value, Control of Learning Beliefs, Self-efficacy for Learning and Performance, and Test Anxiety. Amongst the six motivation scales, Intrinsic Goal Orientation, Task Value, Control of Learning Beliefs and Self-efficacy for Learning and Performance are often regarded as "positive" motivations, as these have been linked to desirable education outcomes. Conversely, Extrinsic Goal Orientation and Test Anxiety are often linked to less desirable education outcomes, and hence are regarded as "negative" motivations.

The learning strategy section also assesses three main constructs (García & Pintrich, 1995): cognitive strategies (i.e., students' ways of processing information from reading materials and lessons), metacognitive strategies (i.e., students' control and regulation of their own thinking processes), and resource management (i.e., students' control and usage of learning resources). Nine learning strategies scales are used to assess these three constructs: Rehearsal, Elaboration, Meta-cognitive Self-regulation, Critical Thinking, Time and Study Environment, Effort Regulation, Peer Learning, and Help Seeking. In all, the MSLQ includes 81 items, to which students respond on a seven-point scale (not at all true of me to very true of me). Rehearsal is generally regarded as a "negative" strategy, as this has been linked to the approach of learning by rote. All other strategies are generally regarded as "positive" strategies, as these are generally linked to higher order thinking. Table 1 presents the overall structure of the MSLQ, along with sample item statements.

Table 1. Structure and sample item statements of the MSLQ

Scale	Construct	Subscale	# Items	Sample Item Statement
Motivation	Value	Intrinsic Goal Orientation	4	The most satisfying thing for me in this course is trying to understand the content as thoroughly as possible.
		Extrinsic Goal Orientation	4	Getting a good grade in this class is the most satisfying thing for me right now.
		Task Value	6	I think the course material in this class is useful for me to learn.
	Expectancy	Control of Learning Beliefs	4	If I don't understand the course material, it is because I didn't try hard enough.
		Self-efficacy of Learning and Performance	8	I'm confident I can understand the most complex material presented by the instructor in this course.
	Affect	Test Anxiety	5	I feel my heart beating fast when I take an exam.
Learning Strategies	Cognitive Strategies	Rehearsal	4	When studying for this class, I read my class notes and the course readings over and over again.
		Elaboration	6	When I study for this course, I write brief summaries of the main ideas from the readings and the concepts from the lectures.
		Organization	4	I make simple charts, diagrams, or tables to help me organize course material.
		Critical Thinking	5	I often find myself questioning things I hear or read in this course to decide if I find them convincing.
	Metacognitive Strategies	Self-regulation	12	When I study for this class, I set goals for myself in order to direct my activities in each study period.
	Resource Management	Time and Study Environment	8	I make good use of my study time for this course.

Strategies	Effort Regulation	4	Even when course materials are dull and uninteresting, I manage to keep working until I finish.
	Peer Learning	3	When studying for this course, I often try to explain the material to a classmate or a friend.
	Help Seeking	4	When I can't understand the material in this course, I ask another student in this class for help.

The first validation study on the MSLQ was conducted by the instrument developers (Pintrich et al., 1991) with a sample of college students in the United States. Confirmatory Factor Analyses (CFAs) performed in this study indicated a five-factor structure for the motivation section, and a nine-factor structure for the learning strategies section. High levels of internal consistency were found for most scales. These findings have since been supported by studies conducted with other samples of college students in the United States (Cho & Summers, 2012), Oman (Alkharusi et al., 2012) and Singapore (Rotgans & Schmidt, 2010).

In comparison to the supporting evidence that has accumulated with regard to college students, little evidence has been published on the validity of the MSLQ at the secondary level. At the time of writing, only two published studies involving secondary students could be located. Erturan Ilker, Arslan and Demirhan (2014) and Karadeniz et al. (2008) both investigated the factor structure of a Turkish version of the MSLQ. The Karadeniz et al. study involved 1114 Turkish students from three primary schools and three secondary schools, while the Erturan Ilker et al. focused only on secondary level students. Both studies indicated that the instrument exhibited a similar factor structure to that found in earlier studies with college students.

Given the comprehensive nature of the MSLQ, this instrument has potential for monitoring the self-regulation levels of secondary students in Singapore. To date, however, no validation studies in this context could be located. Thus, in the present study, the validity of the MSLQ was examined using a Singapore secondary school sample. Two key aspects of validity evidence were examined within the study, based on the guidelines of the 2014 Standards for Educational and Psychological Testing (AERA, APA, & NCME, 2014). The internal structure of the MSLQ was first investigated by replicating the CFAs published in previous validations, and by assessing the internal consistencies of, and inter-correlations between, the MSLQ scales. Correlations with external variables which measure theoretically related constructs were then examined to further evaluate the construct validity of the instrument within this sample.

2. Method

2.1 Sample

Participants were 441 Year 11 students (267 male, 174 female) within the Singapore secondary system (age M = 16.7 years, SD = 0.82). Participants were enrolled in a two-year physics course taught in the English language, which would prepare them for the General Certificate of Education Advanced Level (GCE "A" Level) physics examination, at the time of the study. As such, the MSLQ was administered in this study with specific reference to physics classes, as recommended by Pintrich et al. (1991).

2.2 Validation Instruments

In addition to the Motivated Strategies for Learning Questionnaire (MSLQ), two further instruments were used for validation purposes: the Two-factor Revised Learning Process Questionnaire (R-LPQ-2F; Kember et al., 2004), and a physics achievement test. The R-LPQ-2F includes two main scales (Deep Approach and Surface Approach), each of which includes a motive and strategy subscale (i.e., four subscales in all: Deep Motive, Deep Strategy, Surface Motive and Surface Strategy). Studies on the R-LPQ-2F have indicated that it demonstrates sound psychometric properties (e.g., Phan & Deo, 2007; Socha & Sigler, 2012). Drawing upon the theoretical bases of the two instruments (e.g., Biggs & Tang, 2007; Curran & Bowie, 1998), deep motive scores would be expected to correlate positively with the positive motivation scales in the MSLQ, while deep strategy scores should correlate positively with the positive MSLQ learning strategy scales. Surface motives would be expected to correlate negatively with the positive MSLQ motivation scales, and positively with the two negative MSLQ

motivation scales, while surface strategies should correlate positively with the negative MSLQ strategy scale of Rehearsal. Expected correlations between the MSLQ and R-LPQ-2F are summarized in Tables 2 and 3.

Table 2. Expected correlations between MSLQ motivation scales R-LPQ-2F motive subscales

Subscale	Deep Motive	Surface Motive
1. Intrinsic Goal Orientation	Positive	Negative
2. Extrinsic Goal Orientation	Negative	Positive
3. Task Value	Positive	Negative
4. Control of Learning Beliefs	Positive	Negative
5. Self-efficacy for Learning and Performance	Positive	Negative
6. Test Anxiety	Negative	Positive

Table 3. Expected correlations between MSLQ learning strategy scales and R-LPQ-2F strategy subscales

Subscale	Deep Strategy	Surface Strategy
1. Rehearsal	Negative	Positive
2. Elaboration	Positive	Negative
3. Organization	Positive	Negative
4. Critical Thinking	Positive	Negative
5. Metacognitive Self-regulation	Positive	Negative
6. Time and Study Environment	Positive	Negative
7. Effort Regulation	Positive	Negative
8. Peer Learning	Positive	Negative
9. Help Seeking	Positive	Negative

Given that both motivation and learning strategies are presumed to relate in some way to student achievement, a physics achievement test used within the school was used to provide further information on the validity of the MSLQ. This test is a two-hour pen-and-paper assessment comprising 15 multiple-choice questions and 3 short response questions, that are adapted directly from past year GCE "A" level examinations. The test questions, including the marking scheme, were vetted by the subject coordinator and the head of the physics department, who each had more than ten years of teaching experience, to ensure close alignment to the assessment objectives of the GCE "A" level physics examination (Singapore-Cambridge GCE "A" level physics syllabus 9646, 2014). Scores on the Physics Achievement Test were computed by summing the scores for the multiple-choice questions and the structured questions. The multiple-choice and structured questions were each worth a maximum score of 15 and 65 points, respectively, giving a total maximum score of 80. Expected correlations between the MSLQ scales and physics achievement, based on the underlying theory of the MSLQ, are shown in Tables 4 and 5.

Table 4. Expected correlations between MSLQ motivation scales and physics achievement

Subscale	Physics Achievement Test
1. Intrinsic Goal Orientation	Positive
2. Extrinsic Goal Orientation	Negative
3. Task Value	Positive
4. Control of Learning Beliefs	Positive
5. Self-efficacy for Learning and Performance	Positive
6. Test Anxiety	Negative

Table 5. Expected correlations between MSLQ learning strategy scales and physics achievement

Subscale	Physics Achievement Test
1. Rehearsal	Negative
2. Elaboration	Positive
3. Organization	Positive
4. Critical Thinking	Positive
5. Metacognitive Self-regulation	Positive
6. Time and Study Environment	Positive
7. Effort Regulation	Positive
8. Peer Learning	Positive
9. Help Seeking	Positive

2.3 Procedures

Approval to conduct the research was first obtained from the Human Research Ethics Committee of the University of Western Australia. All procedures used within the study were conducted in compliance with the National Health and Medical Research Council's (2007) Australian Code for the Responsible Conduct of Research. Permission was also granted by the Principal of the participating school. All participating students took the R-LPQ-2F, the MSLQ and the physics achievement test in a single session. They were provided with hardcopies of the questionnaires and optical mark sheets to shade their responses.

3. Results and Discussion

SPSS Version 19 (IBM Corp, 2010) was used to compute all descriptive statistics, Cronbach's α coefficients, and bivariate correlations. LISREL 8.8 (Jöreskog & Sörbom, 2006) was used to conduct all CFAs based on the maximum likelihood estimation method. Prior to the analysis, the scores of all negatively worded items were first reversed. Screening tests for conformity to underlying CFA assumptions were then conducted. These tests generally produced satisfactory results. Inspections of z-scores and Mahalanobis distances indicated no significant univariate or multivariate outliers at the .001 level, and there was no evidence of multicollinearity between variables within the set. Skewness and kurtosis coefficients indicated no significant deviations from normality in the item distributions in terms of kurtosis, though moderate levels of skew across several items were observed. Given this, the PRELIS Normal Scores module in LISREL 8.8 was used to transform the scores prior to conducting the CFAs. Item descriptive statistics for the MSLQ motivation and learning strategies scales appear in Tables 6 and 7, respectively.

3.1 Validity Evidence from Investigating Internal Structure

3.1.1 Confirmatory Factor Analyses

The internal structure of the MSLQ was first investigated by examining its factor structure. CFAs were performed separately for motivation and learning strategies scales, given that the two MSLQ sections are theoretically distinct. For the motivation section, three competing nested models were tested, as presented in Table 8. Model M1 included all item statements from all motivation scales as one factor, given that these all assess facets of learning

motivation. Model M2 tested a three-factor model, with items grouped into the three broad theoretical constructs (value, expectancy and affect) stipulated by Pintrich et al. (1991, 1993). Model M3 tested a six-factor model based on the scales proposed by Pintrich et al. (Intrinsic Goal Orientation; Extrinsic Goal Orientation; Task Value; Control of Learning Beliefs; Self-Efficacy for Learning and Performance; Test Anxiety).

Table 6. Item descriptive statistics for the MSLQ

Motivation Subscale	Item	*M*	*SD*
1. Intrinsic Goal Orientation	1	5.01	1.25
	16	5.52	1.19
	22	5.65	1.06
	24	4.75	1.24
2. Extrinsic Goal Orientation	7	5.23	1.34
	11	5.11	1.27
	13	5.91	1.11
	30	4.73	1.50
3. Task Value	4	5.00	1.20
	10	6.12	0.89
	17	5.13	1.17
	23	5.53	1.01
	26	5.29	1.12
	27	5.69	0.99
4. Control of Learning Beliefs	2	5.99	0.83
	9	5.66	1.15
	18	5.78	.99
	25	5.07	1.24
5. Self-Efficacy for Learning and Performance	5	5.02	1.33
	6	4.38	1.46
	12	5.71	1.03
	15	4.48	1.42
	20	4.94	1.17
	21	5.27	1.22
	29	5.12	1.11
	31	5.04	1.19
6. Test Anxiety	3	3.90	1.70
	8	4.48	1.63
	14	4.31	1.67
	19	3.80	1.49
	28	4.21	1.64

Table 7. Item descriptive statistics for the MSLQ learning strategies scale

Learning Strategies Subscale	Item	*M*	*SD*
1. Rehearsal	39	3.98	1.55
	46	4.78	1.34
	59	5.12	1.25
	72	4.29	1.51
2. Elaboration	53	5.25	1.11
	62	4.77	1.29
	64	5.41	1.03
	67	4.53	1.51
	69	5.45	0.98
	81	4.76	1.21
3. Organisation	32	4.69	1.38
	42	5.60	0.98
	49	4.36	1.50
	63	4.95	1.27
4. Critical Thinking	38	4.94	1.32
	47	5.01	1.21
	51	4.57	1.25
	66	4.82	1.21
	71	4.97	1.22
5. Metacognitive Self-Regulation	33	4.58	1.42
	36	4.14	1.48
	41	5.72	0.92
	44	4.67	1.24
	54	5.01	1.38
	55	4.90	1.34
	56	4.49	1.31
	57	4.62	1.37
	61	4.77	1.27
	76	5.66	0.99
	78	4.74	1.30
	79	5.34	1.20
6. Time and Study Environment	35	5.51	1.17
	43	5.05	1.08
	52	3.68	1.68
	65	4.95	1.60
	70	5.48	1.09
	73	6.19	0.92
	77	3.95	1.38
	80	4.83	1.56

7.	Effort Regulation	37	4.94	1.52
		48	5.13	1.30
		60	5.35	1.30
		74	5.23	1.20
8.	Peer Learning	34	4.50	1.18
		45	4.95	1.15
		50	4.16	1.36
9.	Help Seeking	40	3.94	1.56
		58	5.14	1.31
		68	5.48	1.20
		75	5.62	1.14

Table 8. Models tested for the motivation scales of the MSLQ

Model	Number of Factor(s)	Subscales included in factor(s)
M1	One	All motivation items
M2	Three	Factor 1: Expectancy construct subscales—Intrinsic Goal Orientation, Extrinsic Goal Orientation and Task Value
		Factor 2: Value construct subscales—Control of Learning Beliefs, Self-efficacy of Learning and Performance
		Factor 3: Affect construct subscale—Test Anxiety
M3	Six	Factor 1: Intrinsic Goal Orientation
		Factor 2: Extrinsic Goal Orientation
		Factor 3: Task Value
		Factor 4: Control of Learning Beliefs
		Factor 5: Self-efficacy of Learning and Performance
		Factor 6: Test Anxiety

Another three models were tested for the MSLQ learning strategies section, as presented in Table 9. Model LS1 included the item statements for all the learning strategy subscales as one factor. Model LS2 tested a three-factor model, with items grouped into the three broad theoretical sub-constructs (cognitive, metacognitive and resource management) proposed by Pintrich et al. (1991, 1993). Model LS3 tested a nine-factor model, again based on the subscales of the MSLQ proposed by Pintrich et al. (Rehearsal; Elaboration; Organisation; Critical Thinking; Metacognitive Self-Regulation; Time and Study Environment; Effort Regulation; Peer Learning; and Help Seeking).

Table 9. Models tested for the learning strategies scales of the MSLQ

LS1	One	All learning strategy items
LS2	Three	Factor 1: Cognitive strategies subscales—Rehearsal, Elaboration, Organization, Critical Thinking
		Factor 2: Meta-cognitive strategies subscale—Self-regulation
		Factor 3: Resource management—Time and Study Environment, Effort Regulation, Peer Learning and Help Seeking
LS3	Nine	Factor 1: Rehearsal

	Factor 2: Elaboration
	Factor 3: Organization
	Factor 4: Critical Thinking
	Factor 5: Self-Regulation
	Factor 6: Time and Study Environment
	Factor 7: Effort Regulation
	Factor 8: Peer Learning
	Factor 9: Help Seeking

Two absolute fit indices (the Standardized Root Mean Square Residual, or SRMR, and the relative chi-square value, or χ^2/df) and two relative fit indices (the Comparative Fit Index, or CFI, and the Non-Normed Fit Index, NNFI) were used to assess the fit of each model tested. Accepted cut-offs suggest that in CFAs, good model fit is indicated by a χ^2/df < 5 and SRMR < 0.08, with values greater than 0.90 for the CFI and NNFI (e.g., Browne & Cudeck, 1993; Byrne, 1989; Schumacker & Lomax, 2004). The Goodness of Fit Index (GFI) was not used in this study, based on current recommendations within the field (Sharma, Mukherjee, Kumar, & Dillon, 2005). Differences between nested models within each of the MSLQ sections were evaluated using the chi square difference test (i.e., $\Delta\chi^2$).

The fit indices obtained for each model of the MSLQ in this study are presented in Table 10. As indicated, for the motivation section, the one factor model did not fit the data well. Both the three-factor and the six-factor models met accepted cut-offs for the CFI and NNFI, though the SRMR for the three-factor fell above the recommended cut-off value. All three models differed significantly from one another based on the $\Delta\chi^2$ test. Given these results, it was concluded that the six-factor model provided the best fit to the data.

Table 10. Fit indices alternative models of motivation and learning strategy scales

Model	χ^2	df	χ^2/df	SRMR	CFI	NNFI	$\Delta\chi^2$ Statistics
M1: One-factor (motivation)	2801.19*	434	6.45	0.110	0.89	0.88	M1 & M2: $\Delta\chi^2$ (3) = 854.90, p < 0.05
M2: Three-factor (motivation)	1946.29*	431	4.52	0.092	0.93	0.92	M2 & M3: $\Delta\chi^2$ (12) = 557.18, p < 0.05
M3: Six-factor (motivation)	1389.11*	419	3.35	0.076	0.95	0.95	M1 & M3: $\Delta\chi^2$ (15) = 1412.08, p < 0.05
LS1: One-factor (learning strategies)	5065.30*	1175	4.31	0.092	0.89	0.89	LS1 & LS2: $\Delta\chi^2$ (3) = 166.85, p < 0.05
LS2: Three-factor (learning strategies)	4898.45*	1172	4.18	0.093	0.90	0.89	LS2 & LS3: $\Delta\chi^2$ (33) = 1093.71, p < 0.05
LS3: Nine-factor (learning strategies)	3804.74*	1139	3.34	0.087	0.93	0.92	LS1 & LS3: $\Delta\chi^2$ (36) = 1260.56, p < 0.05

*Significant at α = .001 level.

For the learning strategies scales, again, the one factor model did not fit the data well, though both the CFI and the NNFI fell only marginally below the accepted cut-values. Indeed, the χ^2/df value for the one-factor learning strategies model fell within acceptable parameters. The three-factor model represented a significant improvement on the one-factor based on the $\Delta\chi^2$ test, but this model still fell short of accepted cut-offs based on the SRMR and the NNFI. The nine-factor model was clearly the best-fitting, meeting accepted fit levels for the χ^2/df, CFI, and NNFI, though the SRMR for this model was still somewhat high. Based on these results, the nine-factor was deemed to represent the best-fitting model. These results are comparable to those obtained in previous studies

(e.g., Alkharusi et al., 2012; Pintrich et al., 1991; Karadenzi et al., 2008). The path diagrams (with standardized coefficients) for models M3 and LS3 are presented in Figures 1 and 2, respectively.

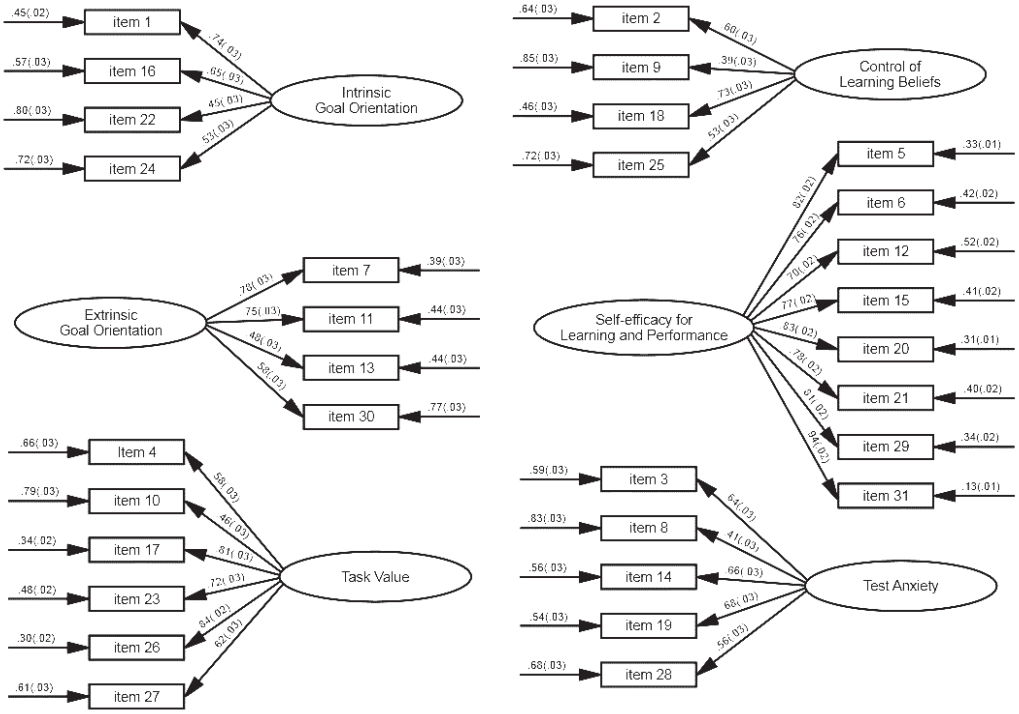


Figure 1. Original six-factor model of the MSLQ motivation subscales

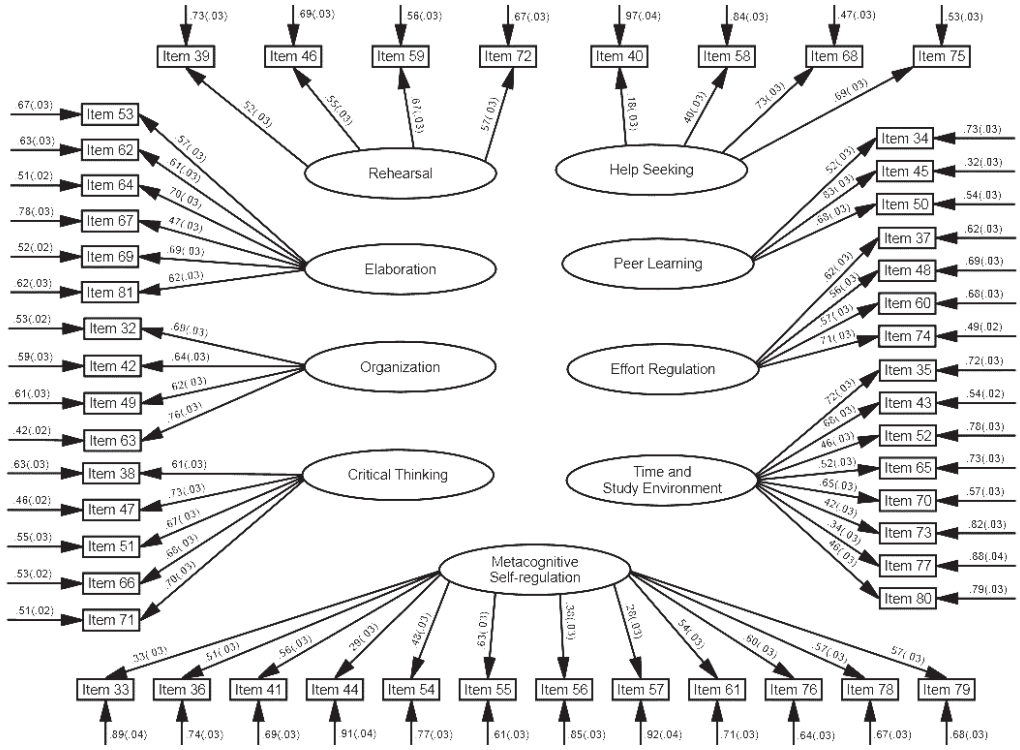


Figure 2. Original nine-factor model of the MSLQ learning strategies subscales

3.1.2 Internal Consistencies

Cronbach's alphas for each of the MSLQ scales are presented in Table 11. Kline (2000) suggested that an alpha coefficient of above .60 represents an acceptable level of internal consistency, with alpha coefficients of above .70 indicating a good level of consistency for low-stakes tests. As indicated, eight of the MSLQ scales had a good level of internal consistency, with six others achieving an acceptable level. Only one scale (Help Seeking) was observed not to meet the acceptable level of 0.6. These results are similar to those obtained by Pintrich et al. (1991). The poor result obtained for the Help Seeking scale is also consistent with previous studies. García and Pintrich (1995) attributed the low internal consistency of Help Seeking to the fact that it relates both to seeking assistance from peers and from teachers. García and Pintrich argued that students might be inclined to seek help from only one of the sources, which would contribute to the lower internal consistency of this scale.

Table 11. Internal consistency of the MSLQ subscales

Scale	Subscale	Cronbach Alpha	
		Current Study	Pintrich et al.'s (1991) Study
Motivation	a) Value Components		
	• Intrinsic Goal Orientation	.76	.74
	• Extrinsic Goal Orientation	.68	.62
	• Task Value	.84	.90
	b) Expectancy Components		
	• Control of Learning Beliefs	.68	.68
	• Self-efficacy for Learning and Performance	.94	.93
	c) Affective Component		
	• Test Anxiety	.69	.80
Learning Strategies	a) Cognitive and metacognitive Strategies		
	• Rehearsal	.69	.69
	• Elaboration	.75	.76
	• Organization	.76	.64
	• Critical Thinking	.80	.80
	• Metacognitive Self-regulation	.79	.79
	b) Resource Management Strategies		
	• Time and Study Environment	.72	.76
	• Effort Regulation	.66	.69
	• Peer Learning	.65	.76
	• Help Seeking	.51	.52

3.1.3 Inter-Correlations between Subscales

The internal structure of the MSLQ was further assessed by examining correlations between the individual subscales. Results are presented in Tables 12 and 13, respectively. The pattern of correlations obtained aligned well with the theoretical basis of the instrument. Intrinsic Goal Orientation, Task Value, Control of Learning Beliefs and Self-efficacy were positively and highly correlated with one other, and all the self-regulated learning strategies factors were also positively correlated with one other. The one subscale that exhibited some unexpected characteristics was the Extrinsic Goal Orientation component. First, this subscale did not correlate negatively with Intrinsic Goal Orientation. This is not aligned with the views of many researchers (e.g., Harter, 1981), who propose the two constructs to be opposing ends of a single continuum. Furthermore, significant positive correlations were found between Extrinsic Goal Orientation and Task Value, Control of Learning Beliefs, and Self-efficacy for

Learning and Performance. This might be deemed counter-intuitive, as many studies conducted in the West have found extrinsic rewards to undermine intrinsic motivation (e.g., Deci, 1971; Kruglanski, Friedman, & Zeevi, 1971; Lepper, Greene, & Nisbett, 1973). The findings of this study are, however, consistent with other studies involving East Asian students. For example, Lin et al. (2003) found that the highest performing Korean students exhibited high levels of intrinsic motivation *as well as* moderate levels of extrinsic motivation. Similar results were also found in a study by Kember, Wong and Leung (1999) on Hong Kong students. Such observations may reflect the Confucian Heritage cultures of these countries, which places great emphasis on hard work and academic success, and in which education is viewed as important not only for the individual, but also, for the family and society (e.g., Biggs, 1998; Salili, 1996).

Table 12. Inter-correlations of the MSLQ motivation scales

Factor		1	2	3	4	5	6
1.	Intrinsic Goal Orientation	–	.05	.67**	.50**	.59**	-.12*
2.	Extrinsic Goal Orientation		–	.24**	.25**	.25**	.27**
3.	Task Value			–	.60**	.58**	.01
4.	Control of Learning Beliefs				–	.48**	-.03
5.	Self-efficacy for Learning and Performance					–	-.24**
6.	Test Anxiety						–

*$p < .05$, **$p < .01$.

Table 13. Inter-correlations of the MSLQ learning strategies scales

Factor		1	2	3	4	5	6	7	8	9
1.	Rehearsal	–	.40**	.58**	.11*	.35**	.26**	.15**	.34**	.25**
2.	Elaboration		–	.59**	.64**	.73**	.44**	.44**	.62**	.34**
3.	Organisation			–	.27**	.57**	.48**	.36**	.47**	.21**
4.	Critical Thinking				–	.64**	.29**	.29**	.45**	.14**
5.	Metacognitive Self-regulation					–	.58**	.56**	.69**	.29**
6.	Time and Study Environment						–	.83**	.61**	.23**
7.	Effort Regulation							–	.78**	.27**
8.	Peer Learning								–	.43**
9.	Help Seeking									–

*$p < .05$, **$p < .01$.

3.2 Validity Evidence Based on Relationships with External Variables

Correlations between the MSLQ scales, R-LPQ-2F motive subscales, and physics achievement test scores are presented in Table 14. With the exception of extrinsic motivation, all obtained correlations aligned well with expectations. The pattern of correlations for Extrinsic Goal Orientation departed somewhat from previous results obtained in Western cultures. Extrinsic Goal Orientation correlated positively (though weakly) with Deep Motive and not with Surface Motive. Again, this may reflect the Confucian Heritage culture of Singapore. Specifically, it may be that even students who adopt extrinsic goal orientations will be motivated to use deep learning strategies, because these strategies are often needed for a high success level.

Table 14. Correlations of the MSLQ motivation scales with R-LPQ-2F

	Factor	Deep Motive	Surface Motive
1.	Intrinsic Goal Orientation	.38**	-.29**
2.	Extrinsic Goal Orientation	.12*	.09
3.	Task Value	.42**	-.29**
4.	Control of Learning Beliefs	.33**	-.16**
5.	Self-efficacy for Learning and Performance	.43**	-.21**
6.	Test Anxiety	-.06	.16**

*$p < .05$, **$p < .01$.

Table 15 presents correlations between MSLQ learning strategies scales, R-LPQ-2F strategy subscales, and achievement. As expected, most of the MSLQ learning strategies scales were positively correlated with Deep Strategy, and negatively or not significantly correlated with Surface Strategy. Help Seeking was an exception, but there were no theoretical grounds for this subscale to exhibit specific correlation patterns with deep or surface learning strategies. Interestingly, it was found that Rehearsal, which is a basic cognitive strategy, was positively correlated *both* with surface *and* with deep learning strategies. While its positive correlation with deep strategies is not aligned with previous studies conducted in Western cultures, this is again consistent with previous studies involving students from Confucian Heritage cultures. Strategies including repetition and memorizing are commonly used by Asian students who engage in deep learning, as a precursor to other strategies (Biggs, 1998).

Table 15. Correlations of the MSLQ learning strategy scales with R-LPQ-2F learning scales and physics achievement test

Factor		Deep Strategy	Surface Strategy
1.	Rehearsal	.11*	.21**
2.	Elaboration	.55**	-.08
3.	Organization	.24**	.05
4.	Critical Thinking	.65**	-.08
5.	Metacognitive Self-regulation	.55**	-.16**
6.	Time and Study Environment	.28**	-.18**
7.	Effort Regulation	.27**	-.26**
8.	Peer Learning	.43**	-.15**
9.	Help Seeking	.06	.06

*$p < .05$, **$p < .01$.

Correlations between the MSLQ scale scores and physics achievement are shown in Table 16. As indicated, again, most of the motivation scales related as expected to achievement, with positive correlations obtained between achievement and Intrinsic Goal Orientation, Task Value, Control of Learning Beliefs, and Self-efficacy for Learning and Performance, and a negative correlation obtained between achievement and Test Anxiety. Extrinsic goal orientation, however, did not correlate significantly with test scores. Relationships between the MSLQ learning strategies scales and achievement (see Table 17) similarly aligned with expectations, though Rehearsal, Organisation, and Help Seeking were not significantly correlated with achievement test scores. Overall, however, these results suggest that the MSLQ scores related to achievement in alignment with the theoretical basis of the instrument.

Table 16. Correlations of the MSLQ motivation scales with R-LPQ-2F motive scales and physics achievement test

Factor	Physics Achievement Test
1. Intrinsic Goal Orientation	.26**
2. Extrinsic Goal Orientation	-.07
3. Task Value	.20**
4. Control of Learning Beliefs	.10*
5. Self-efficacy for Learning and Performance	.27**
6. Test Anxiety	-.10*

*$p < .05$, **$p < .01$.

Table 17. Correlations of the MSLQ learning strategy factors with R-LPQ-2F learning factors and physics achievement test

Factor	Physics Achievement Test
1. Rehearsal	-.05
2. Elaboration	.13**
3. Organization	.05
4. Critical Thinking	.19**
5. Metacognitive Self-regulation	.25**
6. Time and Study Environment	.19**
7. Effort Regulation	.21**
8. Peer Learning	.20**
9. Help Seeking	-.01

*$p < .05$, **$p < .01$.

4. Conclusion

Results of this study indicate that the internal structure of the MSLQ in Singapore high school students is similar to the original factor structure proposed by Pintrich et al. (1991). In this study, a structure comprising six motivation scales and nine learning strategies scales fit the data well. The internal consistencies of the subscales were also generally within acceptable ranges. While the Help Seeking subscale did exhibit a relatively low α coefficient, this aligns with the results obtained in the original validation of the scale. The latter result may signal the need for further refinements of this particular scale. Results of the inter-factor correlations within the MSLQ and correlations with external variables also generally supported the construct validity of the instrument. Overall, these findings suggest that the MSLQ is an appropriate measure of students' learning motivations and strategies in the Singapore secondary context. Future research is needed to determine whether the instrument is suitable for use in other grade levels within this system.

References

AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. Washington, D.C.: Author.

Alkharusi, H. et al. (2012). Psychometric properties of the Motivated Strategies for Learning Questionnaire for Sultan Qaboos. *College Student Journal*, *46*(30), 567-580.

Andreou, E. (2004). Bully/victim problems and their association with Machiavellianism and self-efficacy in Greek primary school children. *British Journal of Educational Psychology*, *74*, 297-309. https://doi.org/10.1348/000709904773839897

Biggs, J. (1998). Assessment and classroom learning: A role for formative assessment? *Assessment in Education: Principles, Policy and Practice*, *5*(1), 103-110. https://doi.org/10.1080/0969595980050106

Biggs, J., & Tang, C. (2007). *Teaching for Quality Learning at University* (3rd ed.). Buckingham: SRHE and Open University Press.

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*, 23-34. https://doi.org/10.1037/0022-0663.93.1.23

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Beverly Hills, CA: Sage.

Byrne, B. M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag. https://doi.org/10.1007/978-1-4613-8885-2

Campbell, M. M. (2001). Motivational strategies, learning strategies and the academic performance of African-American students in a college business environment: A correlational study. *Dissertation Abstracts International*, *62*(2-A), 432.

Cheung, C., Rudowicz, E., Lang, G., Yue, X., & Kwan, A. (2001). Critical thinking among college students: Does family background matter? *College Student Journal*, *35*, 577-597.

Cho, M. H., & Summers, J. (2012). Factor Validity of the Motivated Strategies for learning Questionnaire (MSLQ) in Asynchronous Online Learning Environment (AOLE). *Journal of Interactive Learning Research*, *23*(1), 5-28.

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, *21*, 337-346. https://doi.org/10.1016/j.lindif.2011.03.002

Curran, S., & Bowie, P. (1998). Teaching psychiatry to medical undergraduates. *Advances in Psychiatric Treatment*, *4*, 167-171. https://doi.org/10.1192/apt.4.3.167

Deci, E. L. (1971). The effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*, 105-115. https://doi.org/10.1037/h0030644

Duncan, T. G., & McKeachie, W. C. (2005). The making of the Motivated Strategies for Learning Questionnaire. *Educational Psychologis*, *40*, 117-128. https://doi.org/10.1207/s15326985ep4002_6

Erturan, G. I., Arslan, Y., & Demirhan, G. (2014). A Validity and Reliability Study of the Motivated Strategies for Learning Questionnaire. *Educational Sciences: Theory & Practice*, *14*(3), 9-13.

Eshel, Y., & Kohavi, R. (2003). Perceived classroom control, self-regulated learning strategies, and academic achievement. *Educational Psychology*, *23*(3), 249-260. https://doi.org/10.1080/0144341032000060093

García, T., & Pintrich, P. R. (1995). *Assessing students' motivation and learning strategies: The Motivated Strategies for Learning Questionnaire (MSLQ).* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA: ERIC Document Reproduction Service No. ED 383 770. Available from ERIC database.

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology*, *17*, 300-312. https://doi.org/10.1037/0012-1649.17.3.300

IBM Corp. (2010). *IBM SPSS Statistics for Windows, Version 19.0*. Armonk, NY: IBM Corp.

Jöreskog, K. G., & Sörbom, D. (2006). *LISREL 8.8 for Windows [Computer Software]*. Lincolnwood, IL: Scientific Software International, Inc.

Karadeniz, Ş., Büyüköztürk, Ş., Akgün, Ö. E., Çakmak, E. K., & Demirel, F. (2008). The turkish adaptation study of motivated strategies for learning questionnaire (mslq) for 12-18 year old children: Results of confirmatory factor analysis. *The Turkish Online Journal of Educational Technology*, *7*(4), 108-117.

Kember, D., Biggs, J., & Leung, D. Y. P. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology*, *74*, 261-280. https://doi.org/10.1348/000709904773839879

Kember, D., Wong, A., & Leung, D. Y. (1999). Reconsidering the dimensions of approaches to learning. *British Journal of Educational Psychology*, *69*, 323-343. https://doi.org/10.1348/000709999157752

Kline, R. B. (2005). *Principles and practice of structural equation modelling* (2nd ed.). New York: The Guilford Press.

Kruglanski, A. W., Friedman, I., & Zeevi, G. (1971). The effects of extrinsic incentives on some qualitative aspects of task performance. *Journal of Personality*, *39*, 606-617. https://doi.org/10.1111/j.1467-6494.1971.tb00066.x

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, *28*, 129-137. https://doi.org/10.1037/h0035519

Lin, Y. G., McKeachie, W. J., & Kim, Y. C. (2003). College student intrinsic and/or extrinsic motivation and learning. *Learning and Individual Differences*, *13*(3), 251-258. https://doi.org/10.1016/S1041-6080(02)00092-4

Liu, M. (2003). Enhancing learners' cognitive skills through multimedia design. *Interactive Learning Environments*, *11*(1), 23-39. https://doi.org/10.1076/ilee.11.1.23.13686

McKenzie, K., & Gow, K. (2004). Exploring the first year academic achievement of school leavers and mature-age students through structural equation modelling. *Learning & Individual Differences*, *14*(2), 107-123. https://doi.org/10.1016/j.lindif.2003.10.002

Ministry of Education Singapore. (2010). *21st Century Competencies*. Retrieved from http://www.moe.gov.sg/education/21cc/

National Health and Medical Research Council of Australia. (2007). *Australian Code for the Responsible Conduct of Research*. Canberra, ACT: The National Health and Medical Research Council of Australia.

Neber, H., & Heller, K. A. (2002). Evaluation of a summer-school program for highly gifted secondary-school students: The German pupils academy. *European Journal of Psychological Assessment*, *18*, 214-228. https://doi.org/10.1027//1015-5759.18.3.214

Ostovar, S., & Khayyer, M. (2004). Mathematics self-efficacy and mathemathics problem-solving. Implications of using different self-efficacy forms of assessment. *The Journal of Experimental Education*, *65*, 213-228.

Phan, H. P., & Deo, B. (2007). The revised learning questionnaire: A validation of a Western model of students' study approaches to the South Pacific context using confirmatory factor analysis. *British Journal of Educational Psychology*, *77*, 719-739. https://doi.org/10.1348/000709906X158339

Pintrich, P. R., Smith, D. A., García, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor: Michigan, National Centre for Research to Improve Postsecondary Teaching and Learning.

Pintrich, P. R., Smith, D. A., García, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*, 801-813. https://doi.org/10.1177/0013164493053003024

Rotgans, I. J., & Schmidt, H. G. (2010). The Motivated Strategies for Learning Questionnaire: A measure for students' general motivational beliefs and learning strategies. *Asia-Pacific Education Researcher*, *19*(2), 357-369. https://doi.org/10.3860/taper.v19i2.1603

Sachs, J., Law, Y. K., & Chan, C. K. (2002). An analysis of the relationship between the Motivated Strategies for Learning Questionnaire and the Learning Process Questionnaire. *International Journal of Psychology in the Orient*, *45*(3), 193-203. https://doi.org/10.2117/psysoc.2002.193

Salili, F. (1996). Accepting personal responsibility for learning. In D. A. Watkins, & J. B. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 86-105). Hong Kong: Comparative Education Research Centre, Melbourne: Australia Council for Educational Research.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research*, *58*, 935-943. https://doi.org/10.1016/j.jbusres.2003.10.007

Socha, A., & Sigler, E. A. (2012). Using multidimensional scaling to improve functionality of the Revised Learning Process Questionnaire. *Assessment and Evaluation in Higher Education, 37*(4), 409-425. https://doi.org/10.1080/02602938.2010.536835

Suárez, J. M., González, R., & Valle, A. (2001). Multiple-goal pursuitand its relation to cognitive, self-regulatory, and motivational strategies. *British Journal of Educational Psychology, 71*, 561-572. https://doi.org/10.1348/000709901158677

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Students' Attitudes towards Foreign Languages

Mirjana M. Kovac[1] & Ana Mrsic Zdilar[2]

[1] Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Split, Croatia

[2] Faculty of Civil Engineering, Architecture and Geodesy, University of Split, Split, Croatia

Correspondence: Mirjana M. Kovac, FESB, Rudjera Boskovica 32, 21000, Split, Croatia. Tel: 385-91-430-5715. E-mail: mirjana.kovac@fesb.hr

Received: September 14, 2017 Accepted: October 18, 2017 Online Published: October 24, 2017

doi:10.5539/jedp.v7n2p124 URL: http://doi.org/10.5539/jedp.v7n2p124

Abstract

This paper analyses similarities and differences in attitudes and beliefs towards learning the German and the Italian language. The testing tool used for evaluating the attitudes is a questionnaire composed in accordance with the questionnaires used for examining attitudes towards foreign languages. The obtained results indicate that the students of both languages highly agree with the statement that the knowledge of foreign languages leads to a better social status, more successful socialisation and benefit in one's professional career. In addition, they highly agree with the statement that the teaching materials are miscellaneous, but would nevertheless want the materials to include more elements of the domicile culture. Furthermore, the results also indicate a high level of confidence in speaking both languages and a substantial level of awareness regarding the need to speak both languages. The analysis confirmed a significant difference between the learners of Italian and German; the learners of German perceive the language as more difficult than Italian. However, such subjective evaluation does not have an adverse effect on the attitudes towards the German language.

Keywords: language learning attitudes, language learning motivation, target language

1. Introduction

Numerous studies have been conducted on the students' attitudes, beliefs and motivation concerning language learning. In the past two decades, Horwitz's (1987) *Beliefs About Language Learning Inventory* (BALLI) has been extensively used for various research aiming, for instance, to investigate the links between proficiency and beliefs (Mantle-Bromley, 1995), the influence of culture on beliefs (Horwitz, 1999), the differences between the genders influencing beliefs (Tercanlioglu, 2005), etc.

Students' attitudes towards learning a foreign language are one of the basic factors which influence the learning success. Their significance was recognised by foreign language teachers, and numerous research have emphasised different aspects of their considerably significant role in successful mastering of a foreign language.

Victori and Lockhart (1995, p. 224), in the context of foreign language learning, define beliefs as "general assumptions that students hold about themselves as learners, about factors influencing learning and about the nature of language learning". Horwitz (1988, p. 283) claims that knowing the students' beliefs about a particular language can help teachers to better understand their students "expectations of, commitment to, success in and satisfaction with their classes". This is so, because the teachers, after gaining insight into the student' attitudes towards a particular language, may eventually undertake the necessary steps to alter the negative attitudes.

Karahan (2007) emphasises the crucial role of attitudes towards foreign language learning, and states that a systematic and planned modification of negative attitudes should be performed, because it is essential that the students on their own feel the need to learn a particular foreign language. Choy and Troudi (2006) state that the feelings and emotions of foreign language learners substantially affect their perceptions and attitudes towards that particular foreign language. The increased level of personal motivation and interest in learning of a particular foreign language would most certainly yield better learning results.

Motivation for language learning is a crucial affective factor influencing the learning success of another or a foreign language, because it is directly associated with the frequency of using learning strategies, desire to learn,

and success. Dörnyei and Otto (1998) define motivation as a state in which an individual has been encouraged by a certain stimulus which initiates, directs, coordinates, increases, mitigates and evaluates cognitive and motor processes, whereby the individual manages or fails to successfully realise personal selected desires. Furthermore, motivation is perceived through a time dimension, which means that it is a dynamic and alterable process.

In addition to individual differences, language learning is also affected by cognitive factors referring to intelligence, language talent and speaker's age (Li & Pan, 2009; Kormos & Csizer, 2010). According to Gardner (1985), motivation is composed of a desire to achieve the aimed progress, invested effort and positive attitudes. Positive attitudes in particular help bridge the gap of initial student—teacher interaction in foreign language learning, which may sometimes be stressful to beginners.

Motivation results from attitudes, and Gadner and Lambert (1972) make a difference between integrative and instrumental motivation, which affects learning achievement. Instrumental motivation refers to a student's desire to learn a new language for utilitarian reasons; learning of a new language will reflect on practical values and well-being. On the other hand, integrative motivation aims at better understanding of cultural values of the target language or assimilation into a society or a wider community (Choosri & Intharaksa, 2011). Such people express a sincere and strong desire to learn a different language and its pertaining culture. Instrumental motivation refers to students who wish to learn a foreign language in order to achieve their practical goals, such as employment or earning money. In today's age, there is a strong cosmopolitan tendency not only among the youth, but among all age groups, for various purposes: education, tourism, professional advancement, hence the need to learn other languages besides one's own. Dörneyi (2005), unlike Gardner and Lambert, believes that the integrative and instrumental function are not mutually exclusive, but rather intertwined. He further stresses the importance of factors associated with the learning process, such as the teacher, teaching materials, teaching methods and teamwork.

Noels (2001) mentioned a different dichotomy, i.e., intrinsic motivation resulting from learning interest and personal satisfaction, and extrinsic motivation which refers to instrumental reasons, such as the belief that learning a new language may help with future employment.

Mihaljević Djigunović (1998) also abandons the dichotomic classification into instrumental and integrative motivation, and makes a difference between the utilitarian-communication type of motivation which in fact refers to instrumental orientation (language used for practical and communication purposes), integrative (desire to be part of a social-cultural group which speaks English as a native tongue), and affective type of motivation associated with emotions emerging with speakers whilst speaking a foreign language. In fact, these are aesthetic and emotional components of English language perception. She also presented two demotivating factors which refer to a negative evaluation of the teacher, teaching materials and teaching methods, and difficulties related to English language learning as a consequence of insufficient previous knowledge or lack of talent.

Liu and Zhang (2013) used a substantially large number of Chinese subjects at several universities to examine attitudes and motivation during English language learning. The results indicated that the subjects who demonstrated a high level of integrative motivation spent more time learning, and adopted the target language much faster with respect to the subjects who had a substantially low level of integrative motivation. This may be connected with the tendency of young Chinese to be more involved in the Western culture, and English as a global language represents the best vehicle for such integration.

Balenović (2011) derived from the assumption that adult learners learn the language for professional reasons, and that globalisation processes affect motivation. The findings indicate that the learners demonstrated the highest degree of the affective-cultural motivation, which is stimulated by the affection towards a particular language and the desire to learn more about a foreign culture. It is followed by the professional and instrumental-technological orientation which refers to the need to learn a language for technical or practical reasons. The least evident is the communication-integrative motivation, which refers to learning stimulated by the need to communicate with the speakers of another language.

Mihaljević Djigunović (1998, 2002) conducted extensive research of the role of motivation in the adoption of a foreign language in Croatia. Mihaljević Djigunović and Bagarić (2007) found statistically significant differences in motivation and attitudes towards language by comparing the subjects' age, where older students show lower level of motivation towards German language learning with respect to younger subjects.

In terms of gender differences, Siebert (2003) found that male more than female students believed that the learning of grammar was the most important part of language learning, and using audio-visual materials was of

utmost importance. Bernat and Lloyd (2007) found that gender differed significantly in only two BALLI items, i.e., women were more likely to perceive multilingualism as a feature of intelligence compared to their male colleagues, and women also enjoyed more talking to native speakers compared to their male counterparts.

Didović Baranac et al. (2016) also discovered statistically significant differences in attitudes and beliefs towards German and Spanish learning with respect to gender and age, where female students demonstrated more positive attitudes with respect to their male colleagues, whereas the same differences were not confirmed for the English language. The authors explained the lack of significant differences in English language learning in the context of globalisation processes and the dominant role of English language. In addition, they emphasised that Croatia has become a very attractive tourist destination, hence the need to raise awareness of learning English and other foreign languages. The authors stated that the motivational aspect towards learning German and Spanish increases with age, i.e., that the highest level occurs before high school graduation when the students find jobs or continue their education.

In the research conducted by Mihaljević Djigunović and Bagarić (2007), the attitudes of English language learners are substantially more positive with respect to German language learners. By comparing English and German language learners at the end of elementary school and high school, the authors discovered that the motivation for learning German decreases with the years of learning, and that the attitudes and motivation of subjects are associated with classes, language status and the level of exposure in extracurricular activities.

Cindrić and Narančić Kovač (2005) believe that the status and conditions of learning are different for different languages. English becomes almost a second language and language of international communication, i.e., worldwide *lingua franca*. New conceptualisations regarding the status of English language might explain the differences in attitudes towards other languages which are compared to English. Also, Mihaljević Djigunović and Geld (2002) emphasise the fact that extracurricular exposure to English, compared to German, plays a crucial role regarding attitudes. Bagarić (2001) claims that language exposure stimulates automatic language production. Also, complaints by German learners refer to learning too long and too difficult texts, and that incomprehensible input can play a vital role in language learning and use. It is also worth noting that in addition to extracurricular activities, the students are extensively exposed to English also at home; on TV and Internet, and are thus more comfortable using it even in class.

Dörneyi, Csizer and Nemeth (2006) point out that English and German became the most important languages taught in Hungary after 1989, but, during the past several decades English has become extremely popular as a global lingua franca (McKay, 2003), and languages which had strong regional significance lost a significant amount of importance, such as German with its proximity to Hungary.

In the continuation of the paper, the research methodology, the research results and the conclusion, are presented.

2. Research Aims and Hypotheses

The aim of this research is, based on the questionnaire, to determine similarities and differences in attitudes towards German and Italian in senior-year students of the Catering and Tourist High School. The following hypotheses have been presented:

1) The students perceive German as more difficult with respect to Italian.

The attitudes towards Italian are geographically determined. The learners of German will consider German more difficult than Italian because the Dalmatian dialect is rich with Italian words rooted in speech and culture.

2) The students will have more positive attitudes towards Italian with respect to German.

The attitudes towards the target language depend on the subjective evaluation of the level of difficulty of a specific language. Italian is a more melodic language, and the visual perception of German may be influential in this evaluation, as its words are rather lengthy in comparison to Italian.

3. Methodology

The research was conducted on a group of 35 senior-year students of German and 35 senior-year students of Italian of the Catering and Tourist High School in Split. The testing tool for the students' attitudes is a questionnaire concerning German and Italian. The questionnaire consists of 30 statements, Table 1, and since two languages are concerned, two versions of the questionnaire have been used.

The subjects expressed their level of agreement with a certain statement by circling answers on a 5-degree Likert scale (1 = strongly agree, 5 = strongly disagree). After the collection of data, we proceeded with the statistical

analysis. The questionnaire used for this research was compiled according to the questionnaires used for the purpose of examining attitudes towards a foreign language (Mihaljević Djigunović, 1998; Boonrangsri et al., 2004; Zainol et al., 2012; Burgos & Perez, 2015).

Dörnyei and Taguchi (2010) stated that questionnaires are one of the most frequent and reliable methods for collection data in L2 research, because several types of information can be thus collected; factual information, and information about attitudes and behaviours.

Table 1. Statements to be judged along a five-point Likert scale

No.	Statement
1.	I practice German/Italian whenever I have a chance (watching movies without subtitles, talking to foreign speakers).
2.	I feel that others speak German/Italian better than I do.
3.	I like speaking German/Italian in class.
4.	I feel agitated and uncomfortable whilst speaking German/Italian.
5.	I get uncomfortable speaking German/Italian in front of my colleagues.
6.	I am afraid I will be subjected to ridicule whilst speaking German/Italian.
7.	I do not like to participate in classes of German/Italian on my own initiative. I rather stay silent.
8.	I believe that people who speak foreign languages are more educated and have a higher social standing.
9.	Learning of German/Italian helps me meet new people.
10.	Learning of German/Italian is extremely important because it would benefit my professional career.
11.	People will think better of me and my competences if I speak German/Italian well.
12.	I am forced to learn German/Italian.
13.	I can apply my knowledge of German/Italian in my everyday life (in the summertime, talking to foreigners, friends and similar).
14.	German/Italian is too difficult.
15.	It is fun to learn new phrases and German/Italian grammar.
16.	German/Italian textbooks contain phrases and descriptions which are characteristic of our geographical area (e.g., specific Dalmatian meals and beverages, St. Domnius feast, games, names of Mediterranean plants, descriptions of Dalmatian customs and similar).
17.	Good knowledge of German/Italian positively influences my self-esteem.
18.	I rather learn subjects in Croatian.
19.	I am not particularly interested in German/Italian classes.
20.	Learning of German/Italian is a waste of time.
21.	I intend to invest more time into learning German/Italian.
22.	My teacher always helps me with learning German/Italian.
23.	I speak German/Italian well thanks to my teacher.
24.	Active participation of all students is encouraged in German/Italian classes.

25.	Grammar is more prioritised in German/Italian classes, and verbal communication less.
26.	The importance of teamwork is emphasised in German/Italian classes.
27.	German/Italian classes are rich with various and interesting exercises.
28.	German/Italian teaching materials refer to various fields of interest (politics, sport, fashion, food, communication and similar).
29.	I watch my grammatical accuracy whilst speaking German/Italian.
30.	I wish that German/Italian textbooks contained more elements of the Croatian culture.

4. Research Results

The research results are based on the group of 35 learners of German and 35 learners of Italian. Table 2 illustrates arithmetic means and standard deviations for each of the questions from the questionnaire. Since the arithmetic means for both languages in statements 8, 9, 10, 13, 28 and 30 do not surpass 2.5, it can be concluded that the students agree with said statements regardless of the language. The agreement with statements 8, 9 and 10 implies that the students are aware that the knowledge of foreign languages ensures better social status, a more successful socialisation and benefit in the professional career. The students believe that Italian and German language textbooks contain phrases and descriptions characteristic of a specific area. The agreement with statements 28 and 30 indicates that the students are satisfied with the variety of the content in teaching materials, but they nevertheless would like more elements of their domicile culture in said materials. This may be due to their need to find a similar association between their language and a new language, but also because they would like to know how their culture is verbally reflected in a different language. Vickov (2015) uses the example of the English language to conclude that teaching materials should include the Croatian culture as a starting point, not only because it is necessary to know one's own cultural and historic heritage, but also because of the research results from applied linguistics which indicate that children have a longer and better memory of the meaning of those English words whose content they know from their own culture. Students prefer teaching materials with interesting cultural elements, and the teacher is often seen as a factor causing satisfaction or dissatisfaction with teaching. Mihaljević Djigunović and Bagarić (2007) in their research found that creative and stimulating approaches to teaching German were highly prioritised on the wish list of the learners of German.

Since the arithmetic means for both languages in statements 4, 5, 6 and 20 surpass 3.5, it can be concluded that the students disagree with said statements regardless of the language. The disagreement with statements 4, 5 and 6, which refer to affective components, indicate a high level of confidence whilst speaking a foreign language. The affective component includes the students' emotions, liking or disliking something. Kara (2009) claims that positive attitudes lead to positive behaviour towards studying and De Bot et al. (2005) showed that positive attitudes facilitate language learning. The disagreement with statement 20 indicates a high level of awareness about the need to learn foreign languages.

Table 2. Means and standard deviations for the statements from the questionnaire

Statement No.	German language learners		Italian language learners	
	Mean	**SD**	**Mean**	**SD**
1.	3.400	1.288	3.057	1.282
2.	2.314	0.993	2.657	1.327
3.	2.714	1.296	3.000	1.414
4.	3.686	1.183	3.543	1.197
5.	3.857	1.216	3.771	1.262
6.	4.286	1.100	4.086	1.173
7.	3.514	1.380	3.371	1.330

8.	2.000	1.029	2.086	1.147
9.	1.943	1.187	2.143	1.141
10.	1.657	0.998	1.829	1.043
11.	2.914	1.292	2.914	1.245
12.	3.400	1.376	3.714	1.296
13.	2.143	1.240	2.400	1.168
14.	2.914	1.147	3.543	1.400
15.	3.429	1.220	3.400	1.288
16.	3.314	1.367	2.686	1.471
17.	2.829	1.294	2.686	1.278
18.	2.629	1.190	2.514	1.314
19.	2.800	1.208	3.343	1.349
20.	4.257	0.919	3.800	1.346
21.	3.000	1.306	3.057	1.392
22.	1.857	0.944	2.543	1.039
23.	2.714	0.957	3.029	0.954
24.	2.200	1.183	3.029	1.224
25.	2.914	1.095	2.686	0.900
26.	2.943	0.838	3.486	1.121
27.	2.629	1.003	3.143	1.004
28.	2.200	1.023	2.343	1.259
29.	2.657	1.211	2.800	1.106
30.	2.429	1.119	2.343	1.162

An appropriate statistical test provides the answer to the question of whether the results were obtained from the same population, i.e., if the differences are a matter of coincidence, or they in fact appear because the data were obtained by sampling different populations. Considering that the Likert scale is of ordinal type, and two independent samples are compared, the non-parametric Mann-Whitney test has been used. Test results are presented in Table 3 (0.05 level of significance). It can be noted that statistically significant differences were obtained among students who learn German and those who learn Italian for statements 14, 22, 24 and 26. The existence of statistically significant differences for statement 14 between German and Italian confirms the first hypothesis; i.e., the students perceive German as more difficult with respect to Italian. Other obtained statistically significant differences (statements 22, 24 and 26) refer to the teacher's role in the learning process. The students recognised active participation in class, and the need for teamwork, as well as the contemporary approach to foreign language teaching as an important factor of motivation in language learning and formation of attitudes. Mihaljević Djigunović (2007) investigates the students' satisfaction with teaching. The results of her study point to the conclusion that students favoured teaching based on the communicative approach, where they could express their own ideas. Also, being exposed to authentic materials such as videos or active participation in teamwork increased their motivation for learning the language. In the study conducted by Mihaljević Djigunović and Bagarić (2007), the learners of German thought that textbook topics were outdated and the tasks were not useful enough.

Since no significant differences were obtained for the remaining statements, it can be concluded that the second hypothesis, according to which the students will have more positive attitudes towards Italian, because they are shaped by the belief that German is much more difficult, has not been confirmed. The students perceive German

as more difficult, however such belief is not reflected negatively on other beliefs about the language. Quite the contrary, the students showed positive attitudes towards German regardless of the fact that, judging from their attitudes, it is more difficult with respect to Italian.

Table 3. Results of Mann-Whitney test for each statement

Mann-Whitney test / Statement No.	U	p value	Significant difference
1.	517.0	0.252	No
2.	531.5	0.328	No
3.	534.0	0.349	No
4.	570.0	0.612	No
5.	591.0	0.803	No
6.	533.5	0.353	No
7.	569.0	0.608	No
8.	596.0	0.843	No
9.	526.5	0.287	No
10.	546.0	0.392	No
11.	610.5	0.986	No
12.	534.5	0.356	No
13.	519.5	0.259	No
14.	437.0	0.036	Yes
15.	602.0	0.904	No
16.	458.0	0.066	No
17.	573.5	0.643	No
18.	568.0	0.594	No
19.	473.5	0.095	No
20.	509.5	0.225	No
21.	597.5	0.862	No
22.	386.5	0.006	Yes
23.	503.0	0.160	No
24.	389.5	0.007	Yes
25.	546.0	0.399	No
26.	441.0	0.034	Yes
27.	469.0	0.071	No
28.	595.0	0.834	No
29.	561.5	0.540	No
30.	581.5	0.709	No

5. Conclusion

The results obtained from the conducted research indicate that the learners of the German and Italian language exhibit a strong desire to learn both languages aiming to communicate within the target language community in order to achieve the practical goals, such as employment, earning money and similar.

In addition, learners of both languages showed a high degree of agreement with the statement that teaching materials are miscellaneous, but would nevertheless wish that they contained more elements of the domicile culture. For the learners of Italian, this may be due to the geographical and historic proximity of Italy and its

influence on Dalmatia; for German learners, it may be due to the ever-increasing influence of the ongoing "brain drain" oriented towards German speaking countries which is caused by the social and economic conditions in Croatia. The results also indicate a high level of confidence whilst speaking both languages and a high level of awareness about the need to learn foreign languages.

A statistically significant difference between the learners of German and the learners of Italian has been confirmed; the learners of German perceive the language as more difficult than Italian. However, such subjective evaluation does not have an adverse effect on the attitudes towards German. Other statistically significant differences refer to the teacher's role in the learning process.

The learners of both languages are relatively neutral with regard to the statement that grammar is prioritised in learning, instead of verbal competence. In addition, the agreement with the necessity of grammatical accuracy as a pre-requisite for the development of verbal competence has not been confirmed.

As no significant differences were obtained for the remaining statements, we can conclude that the second hypothesis has not been confirmed, i.e., the subjective perception about the difficulty of the German language did not have an adverse effect on the attitudes towards that language.

It should be emphasised that this research was conducted only on the group of subjects who are graduating from the tourist programme, hence it would be useful to examine whether the obtained results apply to other programmes as well.

References

Bagarić, V. (2001). Jezična svjesnost i učenje engleskog i njemačkog kao stranih jezika. *Strani Jezici*, *30*(3), 107-121.

Balenović, K. (2011). Motivacija odraslih učenika za učenje engleskog jezika u kontekstu globalizacije. *Napredak: Casopis za pedagogijsku teoriju i praksu*, *152*(2), 189-209.

Bernat, E., & Lloyd, M. (2007). Exploring the gender effect on EFL learners' beliefs about language learning. *Australian Journal of Educational and Developmental Psychology*, *7*, 79-91.

Boonrangsri, K., Chuaymankhong, D., Rermyindee, N., & Vongchittpinyo, N. (2004). *The Attitude towards English Language Learning of the Students in Vocational Certificate Level under Curriculum in 2002*. A Case Study of Ayutthaya Technical College studied by Naresuan University, Thailand.

Burgos, E. G., & Perez, S. (2015). Chilean 12th graders' attitudes towards English as a foreign language. *Colomb. Appl. Linguistic Journal*, *17*, 313-324. https://doi.org/10.14483/udistrital.jour.calj.2015.2.a10

Choosri, C., & Intharaksa, U. (2011). Relationship between Motivation and Students' English Learning Achievement: A study of the second-year vocational certificate level Hatyai Technical College Students. *Proceedings-Factors Affecting English Language Teaching and Learning*, 1-15.

Choy, S., & Troudi, S. (2006). An investigation into the changes in perceptions of and attitudes towards learning English in a Malaysian college. *International Journal of Teaching and Learning in Higher Education*, *18*(2), 120-130.

Cindric, I., & Narancic Kovac, S. (2005). English and other foreign languages in Croatia—A needs analysis. *Strani Jezici*, *3-4*, 189-203.

De Bot, K., Lowie, W., & Verspoor, M. (2005). *Second language acquisition: An advanced resource book*. London and New York: Routledge. https://doi.org/10.4324/9780203446416

Didović Baranac, S., Falkoni-Mjehović, D., & Vidak, N. (2016). Ispitivanje stavova prema učenju engleskoga, njemačkoga i španjolskoga jezika kao stranog jezika i jezika struke. *Zbornik sveučilišta u Dubrovniku*, 11-31.

Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, *31*(3), 117-135. https://doi.org/10.1017/S026144480001315X

Dörnyei, Z. (2005). *The Psychology of the Language Learner. Individual Differences in Second Language Acquisition*. London: Lawrence Erlbaum Associates, Inc.

Dörnyei, Z., & Otto, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics (Thames Valley University, London)*, *4*, 43-69.

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in Second Language Acquisition: Construction, Administration, and Processing*. New York: Routledge.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation*. Harlow: Pearson Education.

Dörnyei, Z., Csizér, K., & Németh, N. (2006). *Motivation, language attitudes and globalisation: A Hungarian perspective*. Clevedon: Multilingual Matters.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA: Newbury House.

Horwitz, E. K. (1987). Surveying student beliefs about language learning. In A. Wenden, & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 119-129). Englewood Cliffs, NJ: Prentice-Hall.

Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal*, *72*(3), 283-294. https://doi.org/10.1111/j.1540-4781.1988.tb04190.x

Horwitz, E. K. (1999). Cultural and situational influences on foreign language learners' beliefs about language learning: A review of BALLI studies. *System*, *27*, 557-576. https://doi.org/10.1016/S0346-251X(99)00050-0

Kara, A. (2009). The Effect of a "Learning Theories" Unit on Students' Attitudes Toward Learning. *Australian Journal of Teacher Education*, *34*(3). https://doi.org/10.14221/ajte.2009v34n3.5

Karahan, F. (2007). Language attitudes of Turkish students towards the English language and its use in Turkish context. *Cankaya University Journal of Arts and Sciences*, *1*(7), 73-87.

Kormos, J., & Csizer, K. (2010). A comparison of the foreign language learning motivation of Hungarian dyslexic and non-dyslexic students. *International Journal of Applied Linguistics*, *20*(2), 232-250. https://doi.org/10.1111/j.1473-4192.2009.00247.x

Li, P., & Pan, G. (2009). The relationship between motivation and achievement—A survey of the study motivation of English majors in Qingdao Agricultural University. *English Language Teaching*, *2*(1), 123-128. https://doi.org/10.5539/elt.v2n1p123

Liu, M., & Zhang, X. (2013). An investigation of Chinese university students' foreign language anxiety and English learning motivation. *English Linguistic Research*, *2*(1), 1-13. https://doi.org/10.5430/elr.v2n1p1

Mantle-Bromley, C. (1995). Positive attitudes and realistic beliefs: Links to proficiency. *The Modern Language Journal*, *79*(3), 372-386. https://doi.org/10.1111/j.1540-4781.1995.tb01114.x

McKay, S. L. (2003). Towards an appropriate EIL pedagogy: Reexamining common ELT assumptions. *International Journal of Applied Linguistics*, *13*(1), 1-22. https://doi.org/10.1111/1473-4192.00035

Mihaljević Djigunović, J. (1998). *Uloga afektivnih faktora u učenju stranoga jezika*. Zagreb: Filozofski fakultet u Zagrebu.

Mihaljević Djigunović, J. (2002). *Strah od stranoga jezika*. Zagreb: Naklada Ljevak.

Mihaljević Djigunović, J. (2007). Croatian EFL Learners' Affective Profile, Aspirations and Attitudes to English Classes. *Metodika: Casopis za teoriju i praksu metodikâ u predškolskom odgoju, školskoj i visokoškolskoj izobrazbi*, *8*(14), 115-126.

Mihaljević Djigunović, J., & Bagarić, V. (2007). A comparative study of attitudes and motivation of Croatian learners of English and German. *SRAZ LII*, 259-281.

Mihaljević Djigunović, J., & Geld, R. (2002). English in Croatia today: Opportunities for incidental vocabulary acquisition. *SRAZ*, *47-48*, 335-352.

Noels, K. A. (2001). New Orientations in Language Learning Motivation: Towards a Model of Intrinsic, Extrinsic, and Integrative Orientations and Motivation. In Z. Dörnyei, & R. Schmidt (Eds.), *Motivation Second Language Acquisition* (pp. 43-68). Second Language Teaching and Curriculum Center, University of Havaii, USA.

Scarcella, R. C., & Oxford, R. L. (1992). *The tapestry of language learning: The individual in the communicative classroom*. Boston: Heinle & Heinle.

Siebert, L. L. (2003). Pre-service EFL teachers' beliefs about foreign language learning. *The ORTESOL Journal, 21*, 7-39.

Tercanlioglu, L. (2005). Pre-service EFL teachers' beliefs about foreign language learning and how they relate to gender. *Electronic Journal of Research in Educational Psychology, 53*(1), 145-162.

Vickov, G. (2016). *Hrvatska kultura u učenju stranog jezika*. Zagreb: Hrvatska sveučilišna naklada, Filozofski fakultet u Splitu.

Victori, M., & Lockhart, W. (1995). Enhancing metacognition in self-directed language learning. *System, 23*, 223-234. https://doi.org/10.1016/0346-251X(95)00010-H

Wenden, A. (1991). *Learner Strategies for Learner Autonomy*. London: Prentice Hall.

Zainol, M., Pour-Mohammadi, M., & Alzwari, H. (2012). EFL students' attitudes towards learning English language: The case of Libyan secondary school students. *Asian Social Science, 8*(2), 119-134. https://doi.org/10.5539/ass.v8n2p119

Journal of Educational and Developmental Psychology; Vol. 7, No. 2; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Reviewer Acknowledgements

Journal of Educational and Developmental Psychology wishes to acknowledge the following individuals for their assistance with peer review of manuscripts for this issue. Their help and contributions in maintaining the quality of the journal are greatly appreciated.

Journal of Educational and Developmental Psychology is recruiting reviewers for the journal. If you are interested in becoming a reviewer, we welcome you to join us. Please find the application form and details at http://www.ccsenet.org/reviewer and e-mail the completed application form to jedp@ccsenet.org.

Reviewers for Volume 7, Number 2

Ambrin F Masood, Montana State University Billings, United States

Anjolii Diaz, Ball State University, United States

Elizabeth Fraser Selkirk Hannah, University of Dundee, United Kingdom

Esther S Chang, Soka University of America, United States

Joaquin Linne, University of Buenos Aires/CONICET/University of Lanús, Argentina

Marilyn Espe-Sherwindt, Akron Children's Hospital & Kent State University, USA

Marinella Majorano, University of Verona, Italy

Mercedes Bellido Gonzalez, University of Granada, Spain

Nicole Leach, Mississippi State University, United States

Pathmanathan Moodley, Mpumalanga Department of Education, South Africa

Ryota Nishiyori, National Institutes of Health, United States

Serena Petrocchi, Lugano University, Switzerland

Susan Milne, Campbelltown Hospital, Australia

Call for Manuscripts

Journal of Educational and Developmental Psychology is an open-access, international, double-blind peer-reviewed journal published by the Canadian Center of Science and Education. The journal features a special focus on educational psychology and developmental psychology. The field of educational psychology includes the effectiveness of educational interventions, the psychology of teaching, and the social psychology of schools as organizations and so on. The coverage of developmental psychology includes the following aspects: motor skills and other psycho-physiological processes; cognitive development involving areas such as problem solving, language acquisition; social, personality, and emotional development; and self-concept and identity formation. The journal is available in electronic form in conjunction with its print edition. All articles and issues are available for free download online.

We are seeking submissions for forthcoming issues. All manuscripts should be written in English. Manuscripts from 3000–8000 words in length are preferred. All manuscripts should be prepared in MS-Word format, and submitted online, or sent to: jedp@ccsenet.org

Paper Selection and Publishing Process

a) Upon receipt of a submission, the editor sends an e-mail of confirmation to the submission's author within one to three working days. If you fail to receive this confirmation, your submission e-mail may have been missed.

b) Peer review. We use a double-blind system for peer review; both reviewers' and authors' identities remain anonymous. The paper will be reviewed by at least two experts: one editorial staff member and at least one external reviewer. The review process may take two to three weeks.

c) Notification of the result of review by e-mail.

d) If the submission is accepted, the authors revise paper and pay the publication fee.

e) After publication, the corresponding author will receive two hard copies of the journal, free of charge. If you want to keep more copies, please contact the editor before making an order.

f) A PDF version of the journal is available for download on the journal's website, free of charge.

Requirements and Copyrights

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the authorities responsible where the work was carried out, and that, if accepted, the article will not be published elsewhere in the same form, in English or in any other language, without the written consent of the publisher. The editors reserve the right to edit or otherwise alter all contributions, but authors will receive proofs for approval before publication.

Copyrights for articles are retained by the authors, with first publication rights granted to the journal. The journal/publisher is not responsible for subsequent uses of the work. It is the author's responsibility to bring an infringement action if so desired by the author.

More Information

E-mail: jedp@ccsenet.org

Website: http://jedp. ccsenet.org

Paper Submission Guide: http://submission.ccsenet.org

Recruitment for Reviewers: http://recruitment.ccsenet.org

The journal is peer-reviewed
The journal is open-access to the full text
The journal is included in:

Google Scholar
PKP Open Archives Harvester
SHERPA/RoMEO
Standard Periodical Directory
Ulrich

Journal of Educational and Developmental Psychology
Semiannual

Publisher Canadian Center of Science and Education
Address 1120 Finch Avenue West, Suite 701-309, Toronto, ON., M3J 3H7, Canada
Telephone 1-416-642-2606
Fax 1-416-642-2608
E-mail jedp@ccsenet.org
Website jedp.ccsenet.org

ISSN 1927-0526