

**Impact of intergenerational programs on older adults' psychological well-being:  
A meta-analysis**

by

**Yan Su**

A thesis submitted to the graduate faculty  
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Gerontology

Program of Study Committee:  
Jennifer A. Margrett, Major Professor  
Marcus Crede  
Daniel W. Russell  
Debra M. Sellers

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2017

Copyright © Yan Su, 2017. All rights reserved.

## TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	vi
ABSTRACT	vii
CHAPTER 1. INTRODUCTION	1
Development of Intergenerational Programs	1
Development of Program Evaluation	2
CHAPTER 2. LITERATURE REVIEW	5
Program Outcomes	5
Self-esteem	6
Life satisfaction	7
Depressive symptoms	7
Psychological well-being	8
Health-related quality of life	8
Program Characteristics	8
Types of activity	9
Duration of intergenerational programs	10
Serving types of intergenerational programs	11
Number and ratio of older adults to younger individuals	13
Participant Characteristics	13
Age	14
CHAPTER 3. THEORETICAL FRAMEWORK	15
Social Exchange Theory	15
Elements of social exchange theory applied to intergenerational programs	16
Equity Theory	17
Elements of equity theory applied to intergenerational programs	18
Connecting Variables to Theoretical Frameworks	19
Activity type	19
Duration of intergenerational programs	19
Serving types	20
Ratio of older to younger individuals	20
Study Rationale	20
Research Questions	21
The first aim: Explore program effectiveness	22
The second aim: Examine possible moderators	22

CHAPTER 4. METHODS	23
Design of the Current Study	23
Inclusion and Exclusion Criteria in Searching Literatures	23
Title review criteria	23
Abstract and text review criteria	24
Article Searching and Checking Process	25
Keywords	26
Year of publication	26
Multidisciplinary databases	27
Database search summary	30
Additional search strategies	30
Checking for sample duplication	31
Coding of Studies	31
Data Analysis	32
CHAPTER 5. RESULTS	37
Coding Results	37
Sample Characteristics	37
Description of included studies	38
Results for the First Aim: Effectiveness of the Program	38
Research question one: Overall effect size of the program	39
Results of the Second Aim: Exploration of Possible Moderators	40
Research question two: Program duration	40
Research question three: Serving type, activity type, and ratio of older to younger	41
Research question four: Older and younger participant ages	42
Research question five: Program support and control group	43
Further Exploration: Scatterplots and regressions	43
Effect of Intergenerational Programs on Younger Individuals	44
Meta-analysis Bias	44
CHAPTER 6. DISCUSSION	46
Summary of Findings	46
Implications	47
Program duration	47
Serving type and ratio	48
Older and younger participant ages	49
Program support and control group	51
Logic model	52
Meta-analysis bias	53
Limitations of the Current Study	53
Limitations of Included Studies	55
Recommendations for Future Research	57
Conclusions	63

REFERENCES	64
APPENDIX A. COMPARISON OF LITERATURE SEARCH DATABASES	96
APPENDIX B: ARTICLE CODING MANUAL	98
APPENDIX C: IRB EXEMPTION	104

## LIST OF FIGURES

	Page
Figure 1. Program Characteristics and the Effects on Participants	75
Figure 2. Connecting Social Exchange Theory and Equity Theory with Program Characteristics	76
Figure 3. Article Review Process Using Ten Databases and Other Strategies	77
Figure 4. Funnel Plots of Effect Size against Sample Size for Each Outcome Variable	78
Figure 5. Combination of Variables and Specific Analysis	80
Figure 6. Scatterplots of Effect Sizes against Program and Sample Characteristics	81
Figure 7. Logic model Describing Suggesting Elements Leading to Good Practice of Intergenerational Programs	84

## LIST OF TABLES

	Page
Table 1. Basic Characteristics of All Included Studies	85
Table 2. Intergenerational Program Characteristics	87
Table 3. Intergenerational Program Characteristics and Notes from Authors	88
Table 4. Meta-analysis Results for Older Adults	89
Table 5. Correlations among Dependent and Independent Variables	90
Table 6. Comparing Program Effects among Serving Types, Program Support, and With or Without A Control Group: <i>ANOVAs</i> and Independent <i>t</i> -tests	91
Table 7. Sample Size Weighted Regressions of Participants' Ages and Program Characteristics Predicting Program Effect Size	92
Table 8. Exemplary Program Characteristics Benefiting Older and Younger Participants	93
Table 9. Difficulties in Conducting Meta-analysis for Intergenerational Programs	94
Table 10. Summary of Research Questions and Interpretation	95

## ABSTRACT

A comprehensive evaluation of intergenerational programming (IGP) is needed to identify best practices. In this study, I conducted an IGP evaluation whose first purpose was to explore the effectiveness of IGPs through a meta-analysis of programs reported in the literature between 2000 and 2016. I first examined the effect of IGPs on older adults' depressive symptoms, self-esteem, and life satisfaction separately, and then combined them into a single indicator of psychological well-being. The second aim of the study was to identify possible moderators that might affect the success of IGPs with older adults, including IGP characteristics such as activity type (social activity or personal-related activity), serving type or intended purpose (who is serving whom—older adults' participation benefitting younger participants or vice versa), ratio of older to younger participants, IGP duration (i.e., program length, intervention time per session, number of sessions, and interval between two consecutive sessions), program support (who is facilitating the interaction between younger and older adults), and participant characteristics (i.e., younger and older participants' age). Fifteen studies with 625 older adults were included in the study, and the Hunter and Schmidt (2004) meta-analytic approach was employed to perform the analyses. The results indicated that IGP participation was related to enhanced life satisfaction and self-esteem and reduced depressive symptoms among older adults. Pooling these effects into one construct representing psychological well-being ( $d = 0.37$ , 80% credibility interval =  $[-0.27; 1.00]$ ), I found that IGP was indeed effective for older adults' psychological well-being but not significant. Overall, the effect of IGP was positive; however, the effectiveness was variable, implying the possibility of moderating factors that indeed produce the effectiveness. Due to the limited number of studies, moderator analysis was not conducted; however, the relationships among IGP characteristics, participant characteristics,

and IGP effect size were explored through scatterplots, correlations, analysis of variance (*ANOVA*), independent samples *t*-test, and sample-weighted regression analyses. Of note, younger participants' age ( $r = 0.64$ ,  $p < .05$ ) was significantly related to the higher IGP effectiveness. Additionally, I tried to investigate the effect of IGP on younger participants; however, due to the variety of IGP outcomes I was only able to present the effect sizes of younger participants for individual studies. For future researchers, more investigation regarding IGP effects on younger individuals and more quantitative and comprehensive research utilizing consistent reporting and coding procedures is needed to better understand the overall IGP effectiveness and identify the best practices.



## **CHAPTER 1. INTRODUCTION**

“Society is based on the giving and receiving of resources across the lifespan” (DeVore, Winchell, & Rowe, 2016, p. 216). Both younger individuals and older adults have resources that could be shared with one another such as knowledge, capacity, and skills. When younger individuals share resources with older persons and vice versa, their developmental needs (e.g., older adults’ generativity, integrity, and younger individuals’ identity need) can be met during the process. Intergenerational programming (IGP) is an example when younger individuals and older adults interact with one another in purposeful ways to share their resources.

### **Development of Intergenerational Programs**

IGP is defined as a program that brings together younger and older persons for their mutual benefit (Ward, 1997) through various activities such as entertainment, education, and exercise. One example of mutual benefit would be when older adults have wisdom and experience to share with younger individuals, while at the same time younger individuals can provide companionship to older adults. IGP was originally based on the premise that older adults possess knowledge to be passed down to their grandchildren, creating a process in which both grandparents and grandchildren benefit (Newman et al., 1997). This idea gradually extended to link unrelated, non-familial older adults and younger individuals to help younger individuals achieve better attitudes towards older adults and contribute to the well-being of older adults.

The first documented IGP was implemented in 1963 to provide support for older adults with limited resources and also help children with special needs (Generations United, 2007). Since that time, the development of IGP has gone through tremendous changes. According to an IGP content analysis conducted in 2008, from the 1970s through the 2000s, the number of younger IGP participants expanded from mainly primary school students and preschoolers to

include more college students and fewer preschoolers, although primary school and secondary school students continued to be the dominant group (Jarrott, 2011). In addition, older participants in IGP became more diverse, changing mainly from being community-dwelling older adults without specific health conditions to an increasing proportion of older adults with health challenges (e.g., cognitive impairment, depressed mood, and physical impairment) (Jarrott, 2011).

In addition to the changing nature of IGP participants, we also witnessed the evolution of IGP goals whose development was accomplished in two phases. The first phase addressed generational segregation caused by geographic mobility, family structures, retirement policies, and other issues such as the increasing popularity of using long-term care facilities (Newman, 1997). The second phase focused on solving social issues, including improvement of younger individuals' self-esteem, academic participation and achievement and older adults' mental health, substance use, and level of support (Newman, 1997). Further, an examination of IGP reveals that it has been initiated in a number of countries, including Japan, Sweden, US, Spain, Australia, Canada, and South Africa (Newman & Hatton-Yeo, 2008).

### **Development of Program Evaluation**

Compared with the development of IGP, the advancement of IGP evaluation has occurred relatively slowly and several limitations still exist. There is progress such as an increased application of theories, an increasing number of study articles, and even initiation of *Journal of Intergenerational Relationships* (Jarrott, 2011). However, significant limitations exist related to the scientific and empirical rigor of IGP design (e.g., lack of random assignment and control groups). In addition, evaluation efforts are severely hampered by inconsistent reporting of key components needed to evaluate program effectiveness (e.g., pre/post assessments, measure

reliability, sample characteristics). For example, a content analysis conducted in 2008 that included 127 articles and considered qualitative, quantitative, and mixed-method studies showed that by the 2000s more than 72% of studies reported only benefits of IGP, more than 40% were cross-sectional in design, and only 23.6% represented quantitative research (Jarrott, 2011). From the 1980s through the 2000s, no existing evidence of finding studies reporting both benefits and drawbacks of IGP (Jarrott, 2011). Jarrott (2011) concluded that IGP need improvements in the following fields: multigenerational evaluations, application of theory, larger sample sizes, standardized measures that beyond perceptions and attitudinal change, longitudinal assessments, and rigorous application of analyses.

Another systematic review of IGPs from 1990 to 2012 (Knight, Skouteris, Townsend, & Hooley, 2014) emphasized the importance of reciprocal giving, meaning both younger and older participants can benefit from IGP, and also presented the limitations of the included studies: “lack of empirical rigor in defining and measuring reciprocal giving for both generations, confounding variables such as reminiscence or personal narratives in the shared activity, insufficient use of control groups, and preponderance of cross-sectional design” (p. 275).

It appears that there are noticeable issues in performing IGP evaluations. First, it does not appear that a meta-analysis of IGP has been conducted, although IGP literature reviews have been found that mainly focus on older adult participants with dementia. A previous comprehensive review of the effects of IGP with healthy older adults or a meta-analysis has not been identified. Second, the extant research on IGP evaluation is characterized by contradictory results in terms of benefits to older adults. For example, a Japanese study indicated that IGP was associated with improvement in depressive mood among older adults (Murayama et al., 2015), while other researchers found that IGP had no significant effect on decreasing depressive

symptoms among older adults (Cardona, 2002; Skropeta, Colvin, & Sladen, 2014). A third issue is an absence of consensus regarding recommendations for IGP characteristics and components or an exemplary good practice for other IGP practitioners and researchers to follow. For example, IGPs may include various activities such as singing (Williams, Renehan, Cramer, Lin, & Haralambous, 2012), reading (Isaki & Harmon, 2015), or games (Morita & Kobayashi, 2013). Session durations have varied from 15 minutes (Posada, 2006), to 50 minutes (Hernandez & Gonzalez, 2008), and up to 120 minutes (Gaggioli et al., 2014). Confronted with these diverse traits of IGP, a comprehensive review of IGP such as meta-analysis becomes indispensable. A meta-analysis of IGP could identify optimal ways for connecting younger individuals with older adults thereby contributing to the development of IGP best practices through investigation of the key characteristics and optimal dosages among various IGP traits.

## **CHAPTER 2. LITERATURE REVIEW**

IGP has witnessed tremendous evolution including the program goals, expanding of IGP networking, and the increase of professionals related to IGP, yet the main purpose of IGP has remained the same: to serve younger and/or older participants. To identify the most effective IGP practices to better serve both younger and older participants, an evaluation of IGP should cover all aspects of IGP. As indicated by Ward (1997), evaluation of IGPs are composed of three parts: 1) IGP outcomes; 2) IGP characteristics; 3) and participant characteristics. As evident in the extant literature, evaluation of IGP outcomes has focused on healthy behaviors, attitudes and long-term health outcomes for older adults and on knowledge and attitudes for younger individuals. In this study, evaluation of IGPs focused on components of older adults' psychological well-being. Older adults not suffering from dementia were the target population. The following section provides more a detailed description of IGP outcomes.

### **Program Outcomes**

IGP can facilitate benefits for both younger and older adults. For older adults, IGP improves healthy behaviors and components of well-being such as self-esteem (Gaggioli et al., 2014) and depressive symptoms (Chung, 2009). For younger individuals, IGP may improve their attitudes toward aging and their self-esteem (Sue, 2014). Figure 1 depicts a theoretical model including potential long-term (life satisfaction, happiness, and quality of life) and short-term (loneliness, and knowledge to older or younger individuals) benefits of IGP. The main purpose of this study was to investigate the impact of IGP on older adults' self-esteem, life satisfaction, depressive symptoms, and health-related quality of life (HRQoL). The second purpose was to identify potential moderators of IGP effectiveness. The effect of IGP on younger individuals was also examined.

Concepts of life satisfaction, self-esteem, and depressive symptoms were included as indicators of psychological well-being to assess the effectiveness of IGP on older adults. This choice of variables was made because those concepts were widely used in evaluating IGP (Ward, 1997) and they were intended to be consistent with the large number of articles available for the meta-analysis. As reflected in prior literature, these three variables were correlated with one another and could be grouped together as a single concept. For example, previous studies have found negative associations between depressive symptoms and life satisfaction (e.g., Headey, Kelley, & Wearing, 1993; Sue, 2014; Swami et al., 2007) and positive relation between life satisfaction and self-esteem (e.g., E. Diener & M. Diener, 1995). Self-esteem and depressive symptoms have also been shown to be negatively associated (Battle, 1978; Brown, 1986; Stevens-Ratchford, 1993). Kraut et al. (1998) described psychological well-being as including self-acceptance, purpose in life, and positive relationships (Kraut et al., 1998). Even though the three variables included in this study are not representative of every aspect of psychological well-being, they do comprise key elements of psychological well-being. Refer to Table 1 for detailed outcomes included in the current study.

### **Self-esteem**

In this study, self-esteem has been defined as the evaluation of self-confidence, self-respect, self-worthiness and competence (Branden, 1971), and measurement tools include self-esteem scales such as Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1979) and self-efficacy measurement scales. The reason that the construct of self-efficacy was merged with the construct of self-esteem was because self-efficacy and self-esteem have been found to be highly correlated with each other. For example, Judge, Erez, Bono, and Thoresen (2002) concluded through a meta-analysis that the average correlation of these two variables were 0.7-.8. Further, they

conducted a confirmatory factor analysis combining self-esteem, generalized self-efficacy, neuroticism, and locus of control as indicators of a higher order construct and found that the factor loading of generalized self-efficacy and self-esteem were .77 and .93 individually.

Examples of RSE measurement items are “on the whole, I am satisfied with myself” and “at times I think I am no good at all”. Examples of generalized self-efficacy measurement are “I can always manage to solve difficult problems if I try hard enough” and “it is easy for me to stick to my aims and accomplish my goals”.

### **Life satisfaction**

In this study, life satisfaction refers to the self-assessed reflection of one’s life experience and attitudes about one’s current life situation. The Life Satisfaction Scale (LSS; Salamon & Conte, 1998), Satisfaction With Life Survey (SWLS; Diener, Emmons, Larsen, & Griffin, 1985), and Temporal Satisfaction with Life Scale (TSWL; Pavot, Diener, & Suh, 1998) have been widely used to measure life satisfaction, especially for older adults. For example, the SWLS is a 5-item instrument to measure one’s global satisfaction with life and the two example items are “In most ways my life is close to my ideal” and “I am satisfied with my life” (SWLS; Diener, Emmons, Larsen, & Griffin, 1985).

### **Depressive symptoms**

Depressive symptoms have been defined as an abnormal mood, characterized by sadness, worthlessness, and hopelessness (Anderson, Anderson, & Glanze, 1998). The Geriatric Depression Scale (Yesavage et al., 1983) is a commonly used measurement in assessing depressive mood among older adults. It includes 30 items and the score ranged from 0 to 30, with 0-9 meaning "not depressed", 10-19 referring to "mildly depressed", and 20-30 meaning "severely depressed" (Yesavage et al., 1983). A short-form of The Geriatric Depression Scale

has 15 items ranged from 0 to 15, with a score  $> 5$  points indicative of depression. Two example items would be “Are you basically satisfied with your life? YES / NO” and “Have you dropped many of your activities and interests? YES / NO”.

### **Psychological well-being**

As mentioned above, psychological well-being covers self-acceptance, life purpose, positive relationships, and so on (Kraut et al., 1998). In addition, prior literature suggests that these three variables are highly correlated with one another among older adults. For example, depressive symptoms have been correlated with life satisfaction ( $r = -0.29, p < .01$ ) among college students (Swami et al., 2007), and these two variables have been significantly correlated ( $r = -0.42$ ) among groups ranging from 18-year-olds to 65-year-olds (Headey, Kelley, & Wearing, 1993), with even higher correlation among older adults ( $r = -0.52$ ) (Sue, 2014). Along with life satisfaction and depression, self-esteem was also correlated with life satisfaction ( $r = 0.47$ ) among 13,000 college students in 31 nations (E. Diener & M. Diener, 1995). It therefore seemed proper to group these three concepts into one concept: psychological well-being.

### **Health-related Quality of Life**

HRQoL is defined by the Center for Chronic Disease and Prevention (CDC) as a person's or a group's self-rated health, including physical and mental health, over time (CDC, 2016). Only measurement of Medical Outcomes Study Questionnaire Health Survey was included. It covers physical functioning, role limitations, energy, emotional well-being, social functioning, pain, and general health. One example of the scale would be “In general, would you say your health is” and the responses ranging from 1 to 5 with 1 = poor and 5 = excellent.



### **Program Characteristics**

Ward (1997) stated that, to evaluate the impact of IGP on participants, program information needs to include the number and types of activities, the number of participants in each activity, and the dates of key points. This section presents program information of types of activity, duration of IGP (program length, duration of intervention per session, the interval between durations, and the number of sessions), serving types, the ratio of older to younger individuals, program support (whether the program has a facilitator to assist the interaction between older and younger), and resources of support (who are the program facilitators). The reasons I chose those factors were because, while the extant evaluations of these components are limited, these components are presumed to be critically important to the success of IGPs. For example, lack of support, along with inappropriate activities, was identified as a main reason for the demise of a program (Deutschman, Bruno, & Jarrott, 2003; Salari, 2002). The following section describes these IGP components that are key to program development and may serve as potential moderators.

#### **Types of activity**

While IGP activities are the core of a program (Ward, 1997), there is no consensus in systematically classifying IGP activities, hampering their appropriate choice. One way to classify the types of activity would be to base them on the content or purpose of the activities. Examples of activity content are of music, education, recreational activities (e.g., gardening, bingo, board games), and narrative-based activity (e.g., reminiscence, reading, creative storytelling; Galbraith, Larkin, Moorhouse, & Oomen, 2015). A second method for classifying activities is determining whether or not both younger and older participants can mutually interact with one another. Previous research suggests that activities in which both younger and older individuals can

participate and benefit is more meaningful than other types of activity that are designed for one age group and may not be understood or viewed as relevant by the other age group participant (Galbraith et al, 2015).

In the current study, I chose to study types of activity rather than whether or not participants can mutually interact with one another for the several reasons. First, when considering published work, it is typically impossible to code activity based on whether younger individual and older adults can mutually interact with one another. In addition, capability for younger and older participants to mutually interact depends mainly on the activity type, so measuring the types of activity is more important than measuring whether participants can mutually interact, and identifying an appropriate way to assess the types of activities is really what is needed. Furthermore, because researchers have merely investigated activity types through description rather than a systematic quantitative study, it is necessary to explore this aspect through meta-analysis. In this study, I chose to classify activity type based on activity content. I classified activity type into personalized activity (e.g., telling their own stories) and formatted social activity (e.g., discussion based on a set agenda) because previous research repeatedly suggested that reminiscing was effective to older adults' well-being (Gaggioli et al., 2014).

### **Duration of IGP**

The duration of an event is the time interval over which it happens or exists. I defined the duration of IGP to include the program duration, the duration of intervention per session, number of sessions, and the interval between two consecutive sessions. The duration of the program is defined as the number of weeks over which the program lasts, while the duration of the intervention per session refers to the activity time per session expressed in minutes. The reason I

chose to include the duration of the program was because IGP has various durations that may last three weeks (Gaggioli et al, 2014), two years (Murayama et al, 2015), or seven years (Sakurai et al, 2016). Intervention durations per session vary and may range from 15-30 minutes (Murayama et al., 2015; Sakurai et al., 2016) to 120 minutes (Gaggioli et al., 2014). In terms of intervention duration per session, there is no consensus regarding choice of appropriate time for each session. For example, Low and his colleagues suggested 45 minutes per week was insufficient to produce positive outcomes (Low et al., 2015).

In contrast to the two studies cited above, West and Hutchinson (1992) suggested that the duration of IGP was dependent on the purpose of this program. For example, if the purpose of IGP is to develop a relationship, the duration of IGP should be long enough to satisfy this purpose, while if the IGP is designed for relief from boredom, a shorter term is enough. It seems reasonable to identify an optimal IGP duration if all IGPs have the same stated purpose: to improve older adults' psychological well-being, so in this paper I chose to study one outcome, psychological well-being. Prior work suggests that older adults and younger individuals sometimes established close relationships and became "grandfriends" with this close relationship lasting after this program had ceased (Low et al., 2015). Thus, the optimal duration of IGP is an empirical question. To address this needed, I coded and analyzed four variables related to the duration of IGP: the length of program (weeks), the number of sessions, intervention time per session (minutes), and interval between sessions (weeks).

### **Serving types of IGP**

A beneficial approach to categorizing IGPs would be to categorize them based on the population being served (McCrea & Smith, 1997). As McCrea and Smith indicated, even though

the basic assumption is mutual benefit for younger individuals and older adults, it is usually possible to determine who was serving whom.

Among the various types of IGP, several serving types are possible. McCrea and Smith (1997) provide three types of IGP: older serving younger, younger serving older, and younger and older serving others together. However, Generations United (2007) used younger and older serving each other rather than younger and older serving others together. Younger serving older includes younger individuals visiting older adults in homes or communities to provide older adults knowledge such as technology (Generations United, 2007; McCrea & Smith, 1997). Older serving younger mentoring programs, such as older adults serving as tutors to help younger individuals improve their academic performance (Generations United; McCrea & Smith). Older adults and younger individuals serving each other include older adults and younger individuals cooperating to engage in community service (Generations United, 2007). Older and the young serving others together could mean visiting other communities and providing help (McCrea & Smith, 1997). For this study, I coded the classification of activities based on the four serving types of IGP listed above.

It should be noted that service type can also include a shared-site intergenerational program. Shared-site intergenerational programming is defined by Goyer (2001) as programs where both younger and older simultaneously receive ongoing services at the same setting site. The programs may share resources such as space and also may be operated by the same entity (Jarrott & Bruno, 2007). McCrea and Smith (1997) pointed out that shared-site IGP was an example of “hybrids” IGP (p.82), not easily categorized into the previously described types of IGP, so I excluded shared-site IGP from this study.

### **Number and ratio of older adults to younger individuals**

As mentioned earlier, Ward suggested that the evaluation of the impact of a program on participants requires information on the number of participants in each activity (1997). A review of literature demonstrated that most studies reflect groups of either two or three younger individuals and one older adult or a pairing of one younger individual and one older adult. West and Hutchinson (1992) concluded that a small group works better than a larger group. In their study, three to six pairs of older and younger participants worked better than 15 pairs of younger and older adults (30 participants in total). West and Hutchinson suggested, part of the reason for that large groups did not work as well as small groups is because of the low coordinator to participant ratio (1992). With respect to the older-to-younger ratio, a literature review found the ratio of older to younger varies widely. For example, Cardona (2002) conducted IGP using two or three older adults with one younger individual and Sakurai et al., (2016) carried out an IGP study using six to ten older adults with a class of younger individuals. The ratio of older to younger participants seems as important as the number of total participants with respect to affecting interaction quality between younger individuals and older adults, so it is critical to search for an optimal ratio. Thus, a systematic evaluation of IGPs is warranted to identify the optimal number of participants and ratio of participants.

### **Participant Characteristics**

This section presents participant characteristics, including age and developmental stage/year in school for children, participant sex, and race. I also coded participant information such as the percent of male older adults, the percent of male younger individuals, and the percent of each race category (i.e., White, Black, American Indian/Alaska Native, Asian, or Native Hawaiian/Other Pacific Islander) (Krogstad & Cohn, 2014) for both younger and older

participants. However, due to the incomplete reporting of these information, only age of younger and older participants was illustrated here. More detailed information is given in Table 1.

### **Age**

There is a wide range of available age information for IGP participants. Younger individuals participating in this program range from infants to toddlers, kindergarten (Low et al., 2015), high school, and college students (Chung, 2009), while ages of older participants range from the old (75-85; Breyspraak, Arnold, & Hogan, 2008) to the oldest old (85+; Lee, Camp, & Malone, 2007). In this section, I present ages of younger and older participants as continuous variables and choose as older adults those 60 years old or older and as younger individuals those younger than 30 years. The purpose of this exploration is to examine the relationship between younger participants' age and the total effect magnitude and the relationship between older participants' age and the total effect magnitude.

Additionally, based on Erikson's stages of psychological development (1963), both younger and older adults may have unmet needs, and I expected that when younger and older participants helped to meet each other's needs, the effect size would be the largest. Since I could not test the developmental needs of younger and older participants, I examined the relationship among participants' age as a proxy.

### **CHAPTER 3. THEORETICAL FRAMEWORK**

There has been an increase in applying theories including developmental theories and educational theories in IGP research. For example, the number of articles that has employed theories in evaluating IGP has increased from 0% in the 1970s to 25% in the 1980s and 40% in the 2000s (Jarrott, 2011). Jarrott and Smith (2010) suggest that evaluative and developmental theories in prior literature support the use of IGP, with the critique that these theories don't apply to how IGP works or provides ways for practitioners to evaluate IGP effectiveness.

How to promote IGP, especially how to promote the effectiveness of IGP based on social exchange theory and equity theory, remains an untapped area. In this study, I examined the intergenerational exchange of younger and older participants in IGP using social exchange theory (SET) and equity theory for the following reasons. SET and equity theory are among the most influential theories in the field of interpersonal relationships. SET considers an interpersonal relationship as a process of interactive exchanges, and equity theory emphasizes equity in an interactive process. Together, the application of SET and equity theory to IGP helps illuminate the exchange process between younger and older adults. Since SET also helps us to evaluate the stability and satisfaction of intergenerational exchange between younger and older participants, the application of SET to IGP would facilitate the understanding of the development of relationships between younger and older participants.

#### **Social Exchange Theory**

SET emphasizes benefits and costs in social exchange. In IGP, participants are expected to receive benefits such as feeling good, gaining confidence, and experiencing a sense of achievement in the process. With a goal of maximizing one's own benefit, people continually estimate the rewards and cost in IGP. Therefore, investigation of both benefits and drawbacks is

necessary to identify the program characteristics likely to achieve the optimal effect. Application of SET makes it convenient to examine the cost and benefits in intergenerational exchange. In addition, since previous research reported fairly limited challenges or costs of IGP (Jarrott, 2011), an analysis of both benefits and challenges is necessary. Further, the collection of benefits and challenges of IGP could present evidence-based research results, unveil and address the challenges or costs of IGP which would contribute to the development of IGP.

### **Elements of social exchange theory applied to IGP**

SET is a model that interprets society as a series of interactions between individuals based on personal appraisals of rewards and costs (White, Klein & Martin, 2015). SET views relationships as “results-driven” social behavior with people continuously comparing cost and reward to change their decisions and behaviors to obtain the greatest benefit.

Rewards are anything beneficial to an individual’s interest (White, Klein & Martin, 2015). Rewards from IGP for younger individuals might include improved academic performance, access to more aging knowledge, and reduced social discomfort (Harris & Caporella, 2014), and a more comfortable perception with respect to the idea of aging (Kalisch, Coughlin, Ballard, & Lamson, 2013). Rewards for older adults include providing more meaningful roles during retirement to produce successful aging, improved sense of worth and value, improved mood, and better health (Doll & Bolender, 2010). The value of a reward refers to how much a reward one is receiving means to the person, and this value fluctuates over time and may change with time. For example, younger individuals may value knowledge gain more than older adults who are likely to value emotional benefits (Carstensen, Isaacowitz, & Charles, 1999) as Carstensen’s socioemotional selectivity theory suggests.



Costs represent perceived negative consequences produced from social interaction, which are defined as the reverse of rewards. As noted by Jarrott (2011), the majority of prior research cited the rewards of IGP and few presented costs. The costs for older adults participating in IGP include sharing limited time and space with younger individuals (Ruggiano, 2012). For younger individuals, they stated “doesn’t like to see the residents in pain” and “the needing to respect boundaries and space, and not be noisy” (Biggs & Knox, 2014, p.64).

The process of SET is a cost-benefit analysis process to maximize one’s interest. According to SET we use a cost-benefit analysis to help determine whether we want to establish the relationship in the beginning and decide if we want to continue the relationship using cost-benefit analysis as the relationship develops. This cost-benefit analysis process also applied to the other two components of social exchange theory: comparison level (expectations) and comparison level of alternatives. Here, we summarize the principles in SET which can be applied to IGP (i.e., the analysis of cost and benefit to maximize one’s own interest and comparison level). As shown in Figure 2, points related to IGP success can be obtained from these principles.

### **Equity Theory**

Equity theory, created by Adams (1963), requires partners to abide by reciprocity. Equity theory purports that equity in social comparison is determined through the “input/output ratio” or through “the comparison of a general social reference group” (Cook & Messick, 1983, p.3). An equal exchange is desirable in interpersonal relationships including intergenerational relationships. For example, older adults need generativity while younger individuals need identity formation, which are perfect match (Knight, Skouteris, Townsend, & Hooley, 2014). Equity theory also points out that the perceived fairness is not the same with being satisfaction.

In addition, as indicated by Cook and Messick “the process and factors governing preferences and satisfactions are different from those governing perceptions of fairness” (1983, p.3).

### **Elements of equity theory applied to IGP**

The most satisfactory relationship in IGP is one with equal exchange. In a social network, both older adults and younger individuals can offer human capital for the benefit of others. Human capital refers to the knowledge, skills, and techniques acquired by individuals, while social capital refers to the social network (White, Klein, & Martin, 2015). In social interactions, human capital is extremely important for meeting others’ needs, especially in an equal exchange of human capital with one another. In intergenerational exchanges, as Keil and McClintock suggested the transmission of resources was critical and the important consideration was “the fair allocation of resources among group members” (1983, p.13).

An examination of Erikson’s theory of psychosocial development (1963) suggests that both younger individuals and older adults have plenty to offer each other. Older adults aged 40 to 64 years old are confronted with the need of generativity characterized by care, the care of the next generation, and those 65 and above face the need of integrity featured by wisdom. At the same time, adolescents (13-19) focus on establishing social relationships. Combined equity theory and the basic needs of human development based on Erikson’s stages of psychosocial development, it is noticeable that grouping younger and older adults together could help to gain mutual benefits through meeting one another’s needs in equal distribution of resources. Such reciprocal needs of both generations perfectly connect the younger with the older.

Equity should be defined by the individuals involved, not by society or culture. Factors determining the equality of IGP include relative social resources and minimal interest. Resources refer to anything that could help two individuals achieve their goals. When two partners have

different resources, the one with more resources embodies more power. The principle of minimal interest refers to the fact that the factor that determines the power in a social interaction is dependency on social interaction, and a partner with less interest in establishing a social relationship has more power in the relationship. In a relationship of unequal interest, both partners are not satisfied, and the relationship ends either with an equal exchange or a termination of the relationship.

### **Connecting Variables to Theoretical Frameworks**

#### **Activity type**

The principle of maximizing one's own interest applies to activity types in IGP, and in this study coding of activity type is based on the activity content. Unlike Galbraith, Larkin, Moorhouse and Oomen (2015), who coded activity types as art, music, education, narrative-based, recreation activity, etc., I classified activity type as either personal-related or social activity. I propose that, among various activities, the more personal-related activities such as telling one's own stories could result in greater reward and less costs than social activities based on a set agenda because personal-related activities involve a closer relationship and high quality interaction, guaranteeing that both younger individuals and older adults actively and mutually participate, with older adults more likely to obtain the greatest benefits from the program.

#### **IGP duration**

The duration of IGP expected to maximize one's own interest is determined by the principle of maximizing one's own interest. If the period is minimal, it would most likely not produce the expected benefit, while too long a period would increase the cost of participants in terms of time and energy, so a moderate period of interaction is expected to result in the maximum benefit. Only in programs lasting long enough to establish close interpersonal

relationships can both participants receive positive outcomes. Some research studies have suggested that longer durations of IGP may be more likely to produce positive outcomes, and also have suggested that a 45 minute per week duration was insufficient (Low et al., 2015). In this study, I attempted to identify an optimal period.

### **Serving types**

There are mainly four types of IGP: younger serving older, older serving younger, younger individual and older adults serving each other, and younger and older together serving others, and based on equity theory I hypothesize that the third type can produce the most benefit. For younger serving older or older serving younger, the perceived benefits are unequal so the cost and perceived benefits cannot last for long. When younger and older adults are serving each other, the costs and benefits are more likely to be equal, with older adults benefitting the most from this type of interaction.

### **Ratio of older to younger participants**

Too many younger individuals interacting with one older adult may cause too much pressure and lead to a cost increase because ratio of participants can no longer maximize older adults' interest and produce the expected enhancement of well-being. I therefore searched for the optimal participant ratio.

### **Study Rationale**

Even though an extensive body of study indicates that IGP are beneficial to both younger and older participants, few studies have investigated the components of IGPs. IGP systematic analysis is urgently needed to provide specific IGP information, including activity type, IGP duration, and participant characteristics. Also, as practitioners and policy makers attempt to solve social issues, such as a lack of workforce in the field of gerontology and segregation of younger

and older persons in many societies, an attempt to comprehensively evaluate IGP could help to contribute to solving these issues by improving younger individuals' attitudes towards older adults, enhancing older adults' health conditions, identifying beneficial IGP components, and maximizing IGP impact. This paper, due to its purpose of evaluating IGP effects on older adults and the limitation of extant literature, makes two specific contributions: first, through meta-analysis I assessed the overall average effect size of IGP and present a generalized evaluation of IGP. Second, I explored variables that could possibly contribute to the success of IGP.

To better serve the two aims mentioned earlier: overall effect of IGP and possible moderators, I applied social exchange theory and equity theory to evaluate the development and implementation of IGP and make suggestions for its future development. I first investigated the effect of IGP on older adults' life satisfaction, depressive symptoms, and self-esteem. Then, by combining life satisfaction, depressive symptoms, and self-esteem, I examined the effect of IGP on older adults' overall psychological well-being. The effect of IGP on older adults' HRQoL was studied independently.

### **Research Questions**

The first aim of the study was to evaluate the effectiveness of IGP on older adults, which was addressed by assessing the overall effect size of IGP on older adults' psychological well-being. The second aim was to find possible moderators of IGP that might influence the success of IGP. I hypothesize that for dependent variable psychological well-being, the average effect size would change depending on the type of activities, the duration of IGP, IGP serving types, the ratio of older to younger participants, and the ages of younger and older participants.

**The first aim: Explore program effectiveness**

The first purpose of the study includes one research question: How effective is IGP with respect to older adults' self-esteem, life satisfaction, depressive symptoms, HRQoL, and psychological well-being?

**The second aim: Examine possible moderators**

The second purpose of the study involves multiple research questions. Specifically, these questions incorporate how IGP duration (length of the program, length of intervention per session, number of sessions, and interval between sessions), serving type, ratio of older to younger, age (older and younger individuals), and program support influenced IGP effectiveness on older adults. How do younger participant age and older participant age work together to affect IGP for older adults? How do program support and program length work together to determine the effectiveness of IGP on older adults? The effects of IGP on younger individuals were also examined.

## **CHAPTER 4. METHODS**

### **Design of the Current Study**

The purpose of this study was to conduct a meta-analysis to explore the effects of IGP for older adults. I was not directly interacting with human subjects and information related to human subjects used in this project was unidentifiable, so this project met the criteria for Iowa State University institutional review board (IRB) exemption (Appendix C). The study explored the effect of IGP on older adults' psychological well-being and HRQoL, and for the purposes of this paper, psychological well-being included the following concepts: depressive symptoms, self-esteem, and life satisfaction. The study also investigated the moderator effects of program characteristics (e.g., program length, ratio of older adults to younger individuals, serving type, and activity type) and participants' ages on the effectiveness of the program.

### **Inclusion and Exclusion Criteria in Searching Literatures**

Four steps were needed to conduct this meta-analysis: searching literature, reviewing or checking articles, coding studies, and finally analyzing the coded information. To provide context for the database searches, the review criteria are presented first. This section presents the search criteria for title, abstract, and text, respectively, and is followed by explicit steps of reviewing, coding, and analyzing of studies.

#### **Title review criteria**

To be included in this meta-analysis, the title of articles had to meet one of the following two criteria: either the title contained "intergenerational" or it contained "younger" and "older". If the title indicated older adults' health outcomes without mentioning "intergenerational" or "younger and older", it was still considered appropriate for abstract review. This step was taken to include all possible IGP studies that examined IGP but failed to mention IGP. Articles were

excluded if they focused on interventions to treat specific diseases, and articles containing only “younger,” only “older”, or only “younger’s perception of intergenerational programs” were excluded. Titles containing the words “qualitative research”, “shared-site”, “case study”, and “older adults with dementia” were also excluded.

### **Abstract and text review criteria**

After the title check, the abstracts were reviewed using the same criteria as for the text check. Next, the whole text was checked, and to be eligible for the meta-analysis, the text had to meet the following criteria. 1) The articles had to have been published or presented between January 1<sup>st</sup>, 2000 and December 31<sup>st</sup>, 2016 either in English or Chinese. 2) Articles had to focus on an IGP and each had to be a peer-reviewed research paper, a dissertation, or a thesis. 3) Younger individuals’ ages had to be in a range between 0 and 30 to permit the inclusion of all younger individuals (e.g., infants, school-aged children, college students, and graduate students) and older adults’ ages had to be greater than 60 years of age. More details about age criteria are presented in Appendix B (Coding Manual). 4) While studies should have employed both pretest and posttest assessments, if the study provided enough posttest information that could be used to determine effect size, it was still included. 5) The relationship between the younger and older adults should not be grandchild and grandparent because family membership is not a trait of typical IGP. This criterion was consistent with prior research (Jarrott, 2011; Knight, Skouteris, Townsend, & Hooley, 2014). 6) Older adults in IGP were not described as having cognitive impairment or dementia. Studies describing IGP involving only impaired older adults (e.g., those diagnosed with dementia) were excluded because older adults with dementia have limitations with respect to cognition and functioning (Family Caregiver Alliance, 2016) that would likely limit the effectiveness of IGP. In addition, IGP characteristics intended for use with persons



having cognitive impairment are likely qualitatively different from IGPs designed for persons without specific impairments. 7) Studies that examined the impact of IGP on the targeted outcome variables: depressive symptoms, self-esteem, life satisfaction, or HRQoL were included. Outcomes highly correlated with the targeted outcomes found in previous research studies (i.e., self-efficacy) were also included for the current meta-analysis. 8) Each study had to include evaluation of the effectiveness of IGP using quantitative research methods, including experimental or quasi-experimental group designs. The difference between experimental design and quasi-experimental design was that in quasi-experiments participants were not randomly assigned while participants were randomly assigned in an experiment. Qualitative studies were excluded. 9) Shared-site IGPs were excluded from the analysis because, as Ward (1997) suggested for shared site intergenerational programs, older and younger participants reside together, and these situations constitute a special, qualitatively different case of IGP. 10) The article should present sufficient information to support calculation of effect size. The available sources included both studies that reported effect sizes and studies with data that could be transformed into effect sizes, including raw data, independent samples *t*-test, chi square values or F-test with one degree of freedom, certain types of regression, and exact *p* value and sample sizes for effect sizes. 11) Finally, the study sample was checked to make sure that all studies were unique and without duplication.

### **Article Searching and Checking Process**

Initially, I searched related articles from ten databases, resulting in 27 articles that met the criteria of inclusion. I later conducted a focused search, presented later, on these 27 articles. In addition, since most of these 27 articles came from the *Journal of Intergenerational*

*Relationships* and the *Journal of Educational Gerontology*, I examined all issues of these journals published between 2000 and 2016.

### **Keywords**

A combination of keywords was searched to identify the three key features of IGP: inclusion of two generations, inclusion of older adults, and desired well-being outcomes. To this end, the following keywords were chosen based on prior literature and keywords from related articles: (“Intergeneration\*” OR “trans generation\*” OR “cross age” OR “cross generation\*” OR program\*), (“old\* adult” OR senior\* OR elder\*), and (depress\* OR “self-esteem” OR “life satisfaction” OR “quality of life”)

To determine which search terms were most likely to produce the most relevant articles, I tried a search without the word “program” that yielded a low number of articles. Due to the broad definition of IGP, for example, a particular IGP may be not called IGP but rather a specific intervention such as a program for older adults (Yuen, 2003), so I decided to include the word “program” in the search. This found articles that included programs involving both younger and older adult participants to be part of the search results, regardless of whether they were referred to as “intergenerational” or “program”.

### **Year of publication**

Inclusion criteria for year of publication spanned the interval 2000 to 2016, with the year 2000 chosen as the starting year for the current investigation because Jarrott (2011) had previously conducted a systematic review of IGP that spanned from 1960 to 2000 and contained all studies (qualitative, quantitative, and mixed-methods studies) focused on nonfamilial intergenerational programs (Jarrott, 2011). Furthermore, choice of the period from 2000 through 2016 was believed to yield sufficient articles and a contemporary comparison that would

minimize the time effect or cohort difference. One possible cohort difference would be the goal of IGP that has changed from its initial goal of addressing generation segregation in the 1960s and 1970s to solving social issues in the 1980s and 1990s including younger individuals' low self-esteem, older adults' depression, and inadequate social support, etc. (Newman, 1997).

### **Multidisciplinary databases**

I examined the following ten databases: Scopus, PsycINFO, Educational Resources Information Center (ERIC), Dissertations & Theses Global, PubMed, Sociological Abstract, PsycARTICLE Direct, Web of Science, EBSCO, and Google Scholar. The first search included specific keywords that appeared in abstracts within these ten databases. A second search included the keyword “intergenerational program” that appeared in titles. Due to the varied parameters of the ten databases, different keyword combinations and fields were searched. Specific information about the database searches is presented next (see also Figure 3).

**Scopus, PsycINFO and ERIC.** Combinations of the following terms were searched in Scopus and PsycINFO: (“intergeneration” or “trans generation” or “cross age” or “cross generation” or program\*), (“old\* adult” or senior or elder), and (depress\* or “self-esteem” or “life satisfaction” or “quality of life”). For ERIC, initially the same combination of keywords was searched, and more than 30,000 articles were found. A selection of the first 300 articles were reviewed and no relevant articles were identified, so the keyword “intergenerational program”, instead of the above combination of keywords, was used to search abstracts in ERIC from the year 1998 to the present. Articles were identified if the combination of the terms appeared in the abstracts.

Peer-reviewed journal articles, dissertations, and theses written in English or Chinese published between 2000 and 2016 were retrieved. Next, articles from Scopus, PsycINFO and

ERIC were imported into Mendeley, a free tool for managing references, and after a duplication check, 1,766 articles were available for title review. After reviewing the titles, 170 articles were considered suitable for abstract review. After abstracts were reviewed, the complete text of 53 articles was read. Besides searching in the field of abstract, a keyword of “intergenerational program” was used to search articles with the keyword appears in the title in Scopus, PsycINFO, and ERIC, and then exported into Mendeley. This search produced 109 articles available for abstract review. After abstract review, 53 articles were suitable for text review.

**Dissertations & Theses Global.** Keywords of (“Intergeneration\*” OR “trans generation\*” OR “cross age” OR “cross generation\*” OR program\*), (“old\* adult” OR senior\* OR elder\*), and (depress\* OR “self-esteem” OR “life satisfaction” OR “quality of life”) were used for abstract searching within Dissertations & Theses Global. In addition, a keyword of “intergenerational program” was used for title searching. Articles within abstract searching (n = 608) and title searching (n = 19) in Dissertations & Theses Global were exported. After checking the titles, 219 documents were deemed suitable for abstract review, 27 articles were then prepared for text review.

**PubMed.** Keywords of (“Intergeneration\*” OR “trans generation\*” OR “cross age” OR “cross generation\*” OR program\*), (“old\* adult” OR senior\* OR elder\*), and (depress\* OR “self-esteem” OR “life satisfaction” OR “quality of life”) were used for abstract searching. In addition, a keyword of “intergenerational program” was used for title searching. Articles found by abstract searching (n = 440) and title searching (n = 5) were checked, producing six articles available for abstract review. After reviewing abstracts, three articles were completely read.

**Sociological Abstracts.** Keywords of (“Intergeneration\*” OR “trans generation\*” OR “cross age” OR “cross generation\*” OR program\*), (“old\* adult” OR senior\* OR elder\*), and

(depress\* OR “self-esteem” OR “life satisfaction” OR “quality of life”) were used for abstract searching. A keyword of “intergenerational program” was additionally used for title searching and 196 articles resulted from abstract searching and six articles resulted from title searching. Finally, after title, abstract, and text review, one article from this database met the chosen criteria.

**PsycARTICLES Direct.** A keyword of (“Intergeneration\*” OR “trans generation\*” OR “cross age” OR “cross generation\*” OR program\*), (“old\* adult” OR senior\* OR elder\*), and (depress\* OR “self-esteem” OR “life satisfaction” OR “quality of life”) were used to search within abstracts and a keyword of “intergenerational program” was used to search within titles, producing three articles available for review.

**Web of Science.** A keyword of (“Intergeneration\*” OR “trans generation\*” OR “cross age” OR “cross generation\*” OR program\*), (“old\* adult” OR senior\* OR elder\*), and (depress\* OR “self-esteem” OR “life satisfaction” OR “quality of life”) were used for abstract searching, and the keyword “intergenerational program” was used for title searching. This yielded 5,644 (within abstracts) and 34 (within title) articles. After the title check, 15 articles were available for abstract review, with four advanced to a text review.

**EBSCO.** A keyword of (“Intergeneration\*” OR “trans generation\*” OR “cross age” OR “cross generation\*” OR program\*), (“old\* adult” OR senior\* OR elder\*), and (depress\* OR “self-esteem” OR “life satisfaction” OR “quality of life”) were used for abstract searching, and the keyword “intergenerational program” was used for title searching. From the search of EBSCO, 349 (within abstract) and 79 articles (within title) were eligible for abstract review. After abstract review, two were available for text review.

**Google Scholar.** Due to the limited functions of Google Scholar (for example, long searching terms cannot be recognized) the keyword “intergenerational program”, instead of the collection keywords mentioned above, was searched for the interval 2000 to 2016 with respect to both abstracts and titles, yielding a total number of 1,700 articles. After a title check, 239 articles were considered appropriate for abstract review, and after an abstract review, a total of 70 articles were advanced to a text review.

### **Database search summary**

After merging all the articles available for text review and checking for duplicates, it was found that 27 articles met all inclusion criteria for this meta-analysis. In order to identify all suitable articles, I searched all cited references within the 27 articles as well as all articles citing the 27 articles. Using this tree backward and forward approach, three additional articles were identified as appropriate for the meta-analysis.

### **Additional search strategies**

In addition to searching the ten databases noted above, I also posted a request on an open forum of the Gerontological Society of America requesting that authors forward articles examining the role of intergenerational programs. A similar request was sent to a listserv of aging program directors across the United States. Six professionals responded to the request and three provided articles. Unfortunately, none met our criteria. After noting that most these articles were published in the *Journal of Intergenerational Relationships* and the *Journal of Educational Gerontology*, I also searched all the articles published in these two journals from between 2000 and 2016 and identified two additional articles.

I also examined articles providing a literature review of previous research (i.e., Knight, Skouteris, Townsend & Hooley, 2014) and two met our criteria. Finally, I contacted seven

individual authors asking for the data used in their studies so that I could compute effect sizes, and one author provided data that could be used.

### **Checking for sample duplication**

Next, a check of individual variables used in articles was conducted, using a rationale that, if certain variables were highly correlated with our initial four outcomes, I would keep them, otherwise, I would exclude them. This process produced a new variable, “self-efficacy”, that was included because previous research had shown that self-esteem and self-efficacy were highly correlated and appropriate to merge under one construct. Then, a check of duplicated samples was performed and ultimately, 15 articles were identified, and among them one included only three participants and was therefore not used, so the final result was that 15 articles were used in the meta-analysis.

### **Coding of Studies**

A coding spreadsheet and effect size calculation spreadsheet were developed by Dr. Crede (M. Crede, personal communication, September 26, 2016), and a coding manual was created during the coding process. Based on the calculation spreadsheet, pre-post studies adjusted for control groups and pooled standard deviations were used in computing the effect sizes. The Coding Manual (Appendix B) was created based on the selection criteria. The Coding Manual depicts key elements coded for each article. These elements include basic characteristics of each article (author, year of publication, source of publication, country of the study, and notes taken from each individual article), participant information (mean and standard deviation of ages, the percent of male, and the race distribution in both older and younger participants), program information (ratio of older to younger participants, serving type, program support, resource of program support, design, control group, IGP duration), and effect size related information

(measure of dependent variable, meaning of the measurement, reliability, type and trait of dependent variable, sample size, type of effect size, artificial dichotomization of dependent variable).

I first coded the data independently and then asked colleagues to double-check my work for accuracy, and there were a few minor disagreements between my colleagues' coding and my coding. For example, the coding of the serving type for older adults participating in a class with high school students was originally coded as "serve each other" and the colleague coded it as "older serving younger" (DeMichelis, Ferrari, & Rozin, 2015). These issues were resolved by checking the information in the original article against the coding criteria and finally discussing it with my major professor. Thirty-one effect sizes were ultimately generated from the 15 unique studies. The effect sizes were double-checked and one error was found and corrected.

Each article suitable for the meta-analysis was coded. The coded information includes effect size information, both younger and older participant characteristics and program characteristics. All coded information except for serving type and effect size were double-checked by advanced master and doctoral-level students with methodological training because special knowledge is needed to know the serving type and effect size information. The colleagues were given a model outlining the process of reviewing and checking the coded data, and each colleague was offered one or two articles. The serving type was double-checked by a colleague, and occurrences of discrepant or unclear coding was resolved through discussion with colleagues. The effect sizes were checked twice for accuracy by myself.

### **Data Analysis**

Meta-analytic procedures were employed to synthesize the findings across the many variables of the selected studies. In addition, the meta-analysis technique is helpful in determine



the potential moderators that explain the variability of program effects. Because there were limited reports of reliability of various outcome variables, I chose the method of artifact distribution meta-analysis in Hunter and Schmidt meta-analysis program, implying if a reliability estimate was not reported, then no reliability correction was performed for that variable in the meta-analysis (Hunter & Schmidt, 2004). Standard deviation units or *d* values (Hunter & Schmidt, 2004) was used as the measure of effect size for this study. The estimation of the variation in effect sizes was demonstrated through the upper and lower bounds of the 80% credibility interval. Credibility intervals describes the reliability of effect sizes and the possibility of moderators (Whitener, 1990). If the interval contains zero, then it suggests that the effect does not generalize across situations. That is, there are situations in which the effect is likely to be zero or negative. In other words, the average effect may be positive but there are situations in which the effect is very strong and other situations in which the effect is very weak.

There are two models of meta-analyses: fixed-effect and random-effect models. While random-effect models can be generalized to a population of all potential studies investigating IGP effect, estimates from fixed-effect models cannot be generalized (Hunter & Schmidt, 2004). In the current study, I chose the random-effect model so that results obtained from the current meta-analysis can be used to assess IGP effect of all potential studies investigating targeted IGP outcomes.

By using meta-analytic approach, research findings were integrated to identify the overall effect size of IGP. The average effect size can help researchers understand the effect of the program on older adults and direct the future development of conducting research in this field. Cohen's *d* measure of effect size was chosen for the current meta-analysis and interpreted as

follows:  $\sim .20$  was small,  $\sim .50$  was moderate, and anything greater than  $.80$  was a large effect (Cohen, 1988).

In addition, data were transformed into SPSS version 24 for analyses. Missing information comes from the unclear description of variables. For example, when older adults' age was above 65 rather than an age range or a mean age, it was then considered as missing and when the interventions were composed of activities without detailed time, it was then taken as missing information. All analyses involving variables with missing values were conducted using a pairwise deletion approach, meaning that not including the participant when the variable has a missing value, but include the participant when analyzing variables without missing values (Cottrell, & Newman, 2015). Pairwise analyses take into account all available data, however, the disadvantage is that each analysis is based on a different sample sizes (Cottrell, & Newman, 2015).

The independent variables were program information (i.e., the length of the program, the length of the intervention per session, interval between two sessions, activity type, serving type, ratio of older to younger participants) and participant information (i.e., age) meeting the criteria of inclusion. Four variables: depressive symptoms, self-esteem, life satisfaction, and HRQoL were used as outcome variables, and I chose them because I am interested in psychological well-being of older adults. While there are other health-associated outcomes such as generativity, self-efficacy, and self-confident, only self-efficacy was included in our analysis because prior literature suggested that self-efficacy and self-esteem were highly correlated and suitable to be used as indicators of a common construct (i.e., Judge, Erez, Bono, & Thoresen, 2002).

For the dependent variable: psychological well-being, different types of analyses were conducted including scatterplots, correlation, ANOVA, independent samples *t*-test, and sample-

weighted linear regression. With respect to correlation, I conducted correlation analyses among the variables: overall effect size, participant characteristics (i.e., younger participants' age, older adults' age), program characteristics (i.e., program length, interval between sessions, number of sessions, intervention per session, ratio of older to younger). I performed an analysis of variance to examine the effect of serving types on IGP and two independent samples *t*-tests to investigate the different effect sizes studies with or without program support and with or without control group. Multiple regression analysis was conducted to explore the effect of multiple predictors (program length and program support; age of younger participants and age of older participants) on the effectiveness of IGP for older adults.

Finally, a funnel plot was created to explore the bias in the current meta-analysis. One way to examine bias in meta-analysis is through creation of a funnel plot which “plots of effect estimates against sample size” (Egger, Smith, Schneider, & Minder, 1997, *p.*629). The funnel plot is based on the premise that that precision of estimating the treatment effect will enhance when the sample size increases. Egger suggested that a visual examination of funnel plots was helpful to examine the presence of bias in meta-analyses. Effect sizes from studies utilizing small samples will be located widely and loosely at the bottom of the graph while the effect sizes utilizing larger samples will gather more tightly, thereby more closely representing the true effect size (Egger, Smith, Schneider, & Minder, 1997). Egger et al describe an asymmetrical and skewed plot as an indication of publication and/or other biases (Egger, Smith, Schneider, & Minder, 1997). “The sources of asymmetry of the funnel plot include publication bias, choice of effect measure, poor methodological design of small studies, inadequate analysis, English language bias, and citation bias” (Egger, Smith, Schneider, & Minder, 1997, *p.*633). In addition,

when meta-analyses utilized a small sample size, the interpretation of bias in meta-analysis should be cautious (Egger, Smith, Schneider, & Minder, 1997).

## **CHAPTER 5. RESULTS**

This section outlines the results of coding, meta-analysis, correlation, analysis of variance (*ANOVA*), independent samples *t*-test, and regressions. Overall, the findings demonstrated the beneficial effects of IGP on older adults' psychological well-being and also revealed major variance for the effects suggesting the potential for moderating influences.

### **Coding Results**

After title, abstract, and text reviews, 15 articles were suitable for coding. One of the articles studied two separate samples: older adults with and without depressed mood determined by the cut-off point of five on the Geriatric Depression Scale (GDS; Kamei et al., 2011). GDS-15 had a score of 0-15 with higher score meaning more depressive symptoms and a score  $\geq 5$  was considered as depressive symptoms. This study included both of the samples: older adults without depressed mood ( $n=9$ ,  $GDS=2.8$ ) and older adults with depressed mood ( $n=5$ ,  $GDS=8.2$ ) (Kamei et al., 2011). Since the participants in two samples experienced the same IGP characteristics, the two samples were combined as one sample through combining the mean and standard deviation (pooled standard deviation) of the two samples. Therefore, a total of 15 unique studies were included. Based on the criteria described in the Methods section (e.g., targeted outcomes and sufficient data to get the effect size), data including means and standard deviations were extracted from the articles and coded.

### **Sample Characteristics**

Sample characteristics showed a total of 31 effect sizes computed from the 15 unique studies included in the current study. Among these, 21 effect sizes reflected older adults' outcomes and 10 effect sizes reflected younger participants' outcomes. Among the 21 effect sizes for older adults, all four outcomes were represented, although at different frequencies:

depressive symptoms ( $n=7$ ), self-esteem ( $n=5$ ), self-efficacy ( $n=2$ ), life satisfaction ( $n=6$ ), and HRQoL ( $n=1$ ).

### **Description of included studies**

There were various sources of IGP; 46.7% IGP were conducted in the US and 20% in Japan, with other countries that included Canada, Israel, Italy, and Spain. Among the 15 articles, six were theses or dissertations and nine were published articles. Among the nine published articles, five were published in the *Journal of Educational Gerontology*.

Nine of the 15 studies had program lengths of no longer than three months, and 13 of the 15 unique studies had a program length shorter than one year. All reported design methods of the 15 studies were pretest and posttest assessments. In terms of serving types, ten studies were coded as serving each other, two were younger serving older, and four articles were coded as older serving younger. Participants from 11 of the 15 studies met once a week, while other participants met two or three times a week ( $n = 3$ ) and once every two or three weeks ( $n = 2$ ). Nine out of the 15 studies had program support. Additional details are outlined in Table 2.

### **Results for the First Aim: Effectiveness of IGP**

This study embraced two aims: to examine the effectiveness of IGP and to investigate and identify potential moderators. There were five research questions related to the two aims. The first aim was composed of one research question: How effective were IGPs for older adults? The second aim consisted of four research questions (research questions two through five). This section presents the results with respect to these two aims, followed by further exploration, younger individuals' effect size, and investigation of meta-analysis bias.

### **Research question one: Overall effect size of the program**

There was only one study investigating HRQoL, so this variable was not included in the meta-analysis or other analyses. Table 4 presents detailed results for the first aim of the current study. I first examined the degree of IGP effectiveness related to older adults' self-esteem, life satisfaction, depressive symptoms, and psychological well-being, using a random effects model. A summary of the effect sizes is outlined in Table 4.

First, I found that the mean true effect size of IGP related to older adults' depressive symptoms was -0.26, suggesting that IGP was associated with fewer depressive symptoms. However, the standard deviation of delta was 0.51, indicating that the true effect size was variable. In addition, the 80% credibility interval was between -0.92 and 0.41, suggesting that the true effect size might possibly be zero. The credibility interval is used in identifying the existence of moderators (Hunter & Schmidt, 2004).

For older adults' life satisfaction, the average true effect size was 0.11, indicating that IGP was associated with greater life satisfaction. A standard deviation of delta of 0.30 and an 80% credibility interval between -0.27 and 0.49, implying the possible moderators existed for the effect of IGP on older adults

The average effect size of IGP on older adults' self-esteem was 0.20, indicating that IGP was associated with enhanced self-esteem, a standard deviation of delta of 0.25 and an 80% credibility interval between -0.12 and 0.52. Since the credibility interval included zero, it indicated that possible moderators existed for the effect of IGP on older adults' self-esteem.

Finally, life satisfaction, self-esteem, and depressive symptoms were combined as indicators of psychological well-being. Analysis indicated a true effect size of 0.37, implying that IGP was effective with respect to older adults' overall psychological well-being. However,

the standard deviation of delta was 0.50 with an 80% credibility interval of between -0.27 and 1.00, signifying the presence of zero effect size and moderators.

### **Results of the Second Aim: Exploration of Possible Moderators**

The second aim was comprised of four research questions (research questions two to five) that explored the effect of IGP duration, serving type, ratio of older to younger, participant's ages, and with or without a program support on IGP effectiveness. In this study, psychological well-being was considered as the representative of IGP effect on older adults since it incorporated three variables: depressive symptoms, life satisfaction, and self-esteem. The value of depressive symptoms was reversed so higher score refers to favorable effects. Then the values of depressive symptoms, life satisfaction, and self-esteem were averaged to get the effect size of psychological well-being. Q-Q plot of effect sizes showed an approximately normal distribution as all points were close or near the Q-Q line. All following analyses focused on the associations between older adults' psychological well-being, participant characteristics, and IGP characteristics. Figure 5 illustrates the combination of variables and specific analyses.

#### **Research question two: Program duration**

This research question explored whether IGP duration affected IGP effectiveness on older adults. As mentioned in the Literature Review section, four variables were designated by IGP duration: program length, intervention length per session, number of sessions, and the interval between two consecutive sessions. The definitions of these four variables are provided in the Coding Manual (Appendix B). Four correlations with each of the duration variables and the effect size of psychological well-being were conducted (see Table 5). None of the correlations were significant.



Scatterplots (Figure 6 Panel A, B, C, & D) were also employed to investigate the relationship between IGP duration and the effect size, and the following observations are based on the scatterplots. In terms of interval between sessions, the majority of studies (10 out of 14) indicated a meeting frequency of once a week. There were three studies stating that younger and older participants met less frequently than once a week. With respect to program length, 14 out of the 15 studies lasted no longer than one year. Among the 14 studies, the effect varied significantly among IGP that lasted no longer than six months while IGP effects were relatively high for programs between six months and 12 months. Regarding intervention time per session, 12 studies reported the exact intervention time. Among which, a short intervention time such as 15 minutes and 30 minutes were too short to have a positive effect on older adults' psychological well-being. In terms of the number of sessions, the scatterplot was similar to that of the program length showing that no existing evidence of an association between IGP effect and the number of sessions.

### **Research question three: Serving type, activity type, and ratio of older to younger adults**

This question was to investigate how serving types and ratio of older to younger affected the effectiveness of IGP on older adults. An *ANOVA* was conducted to examine the differences among the various serving types. Analyses indicated no significant difference in serving types on older adults' psychological well-being (see Table 6). However, the mean effect size for younger and older serving each other was 0.46 ( $SD = .71$ ), while the mean effect size for younger serving older was 0.37 ( $SD = 1.18$ ), and the mean effect size for older serving younger was 0.15 ( $SD = 0.30$ ). This non-significance could be due to the unequal distribution of the serving types and the limited number of studies involving younger serving older ( $n = 2$ ) and older serving younger ( $n = 4$ ) compared with the number of studies in which younger and older served each other ( $n = 10$ ).

In terms of activity type, since only one study was coded as personal-related activity while 15 other studies were coded as social activity, no analyses were conducted. Next, I examined how the ratio of older to younger participants in a group influenced the effectiveness of IGP on older adults. No significant value was found in the correlation analysis perhaps it was because the variable of ratio was not continuous. As shown in the scatterplot (see Figure 6 Panel E), five studies reported a ratio of one older adult with one younger individual; there were no significant associations between IGP effect size and the ratio of older to younger.

#### **Research question four: Older and younger participant ages**

This research question explored how the age of older and younger participants influenced IGP effect on older adults. A correlation analysis between the age of older participants and the effect size was not significant but showed a trend of negative correlation. The scatterplot showed an effect size difference among 60-70, 70-80 and 80-90-year-olds (see Figure 6 Panel F). To further investigate, I calculated the mean effect size of the three age groups and noted that, for the 60-70-year-olds ( $n = 4$ ) the mean effect size was 0.05 ( $SD = 0.10$ ), for the 70-80-year-olds ( $n = 5$ ) the mean effect size was 0.92 ( $SD = 0.55$ ), and for the 80-90-year-olds ( $n = 5$ ) the mean effect size was 0.07 ( $SD = 0.76$ ).

In terms of how the age of younger participants influenced IGP effect on older adults, a significant and positive correlation between the age of younger individuals and the effect size was found. As shown in the correlation table (Table 5),  $r = 0.65$ ,  $p < 0.01$ , indicating that younger individuals' age was highly and positively correlated with the program effectiveness. A further examination of the scatterplot (Figure 6 Panel G) suggests that IGP was most effective to college students; however, results were not able to support the effectiveness of IGP with respect to preschool students.

### **Research question five: Program support and control group**

First, to examine how program support affected IGP impact on older adults, an independent samples *t*-test about the effect size and program support was conducted. The result was nonsignificant  $t(13) = -2.04, p = 0.06$  yet, as shown, the mean effect size for IGP without program support was 0.03 ( $SD = 0.58$ ) and the mean effect size for IGP with program support was 0.65 ( $SD = 0.59$ ). Next, an independent samples *t*-test about the effect size and control group was run. Results suggest that studies with a control group had lower effect sizes ( $n = 10, d = 0.19, SD = 0.62$ ) than studies without a control group ( $n = 5, d = 0.71, SD = 0.61$ ).

### **Further Exploration: Scatterplots and Regressions**

In addition, with younger individuals' age and older adult age as the two y-axis and effect size as the x-axis, I examined the relationship between younger individual age, older adult age, and the effect size. As shown in Figure 6 Panel H, when the oldest older adults and the youngest younger individuals interacted with one another, the program was less effective. For the oldest older adults, the best choice of younger partners was school-age children (8-11 years old). The most effective cooperation was between 70-80-year-olds and college-aged younger individuals (17-19-year old). I also noted that in the 15 studies, 60-70-year-olds mostly interacted with school-age children and few interacted with preschool children or college students.

Finally, I investigated how younger participant age and older adult age worked together to influence IGP effectiveness with respect to older adults, and how program length and program support worked together to affect IGP impact on older adults. The results showed that younger participant age and older participant age accounted for 61% of the explained variance (Adjusted  $R^2 = .61, F(2, 11) = 11.22, p < .01$ ). The results indicated that younger participant age significantly predicted IGP effectiveness ( $\beta = .81, p < .001$ ). Program support and program length

accounted for 16% of the explained variance (Adjusted  $R^2 = 16\%$ ,  $F(2,12) = 2.35$ ,  $p = .14$ ].

Neither program support nor program length significantly predicted IGP effectiveness. Refer to Table 7 for more details.

### **The Effect of IGP on Younger Individuals**

In terms of younger individuals, ten effect sizes were coded including semantic differential scale ( $n = 2$ ), stereotype or aging attitudes related to older adults ( $n = 3$ ), self-esteem and self-efficacy ( $n = 4$ ) and life satisfaction ( $n = 1$ ). Due to the variability of the outcomes, the ten effect sizes from the six studies were not integrated to generate an average effect size. Table 1 outlines the specific measurement scales for younger individuals and Table 8 provides the effects of IGP on younger individuals (a list of the effect sizes). The values of effect sizes among younger individuals were reversed so higher values signified favorable IGP effectiveness. The correlation between the effect sizes of younger individuals and older adults were also conducted and the result showed a correlation of 0.32.

### **Meta-analysis Bias**

Funnel plot were created to estimate the presence of bias in the meta-analysis. As outlined in previous literature, a graph depicting skewed and asymmetrical effect sizes is believed to indicate the presence of bias (Egger, Smith, Schneider, & Minder, 1997). The graph in the current study was not symmetrical, implying the possibility of bias. For comparison, I attached a symmetrical graph of a funnel plot for the reference of readers (see Figure 4e). All plots contained both studies with small sample sizes and large sample sizes, however, all of the four funnel plots included more positive effect sizes compared with negative effect sizes. A closer examination of the four plots showed that meta-analysis in self-esteem and life satisfaction included were more biased compared with studies in depressive symptoms. The

overall graph of psychological well-being implied the presence of bias. As Egger et al. (1997) suggested that when the sample size was very small, researchers need to be cautious in representing the presence of bias in meta-analyses, this studies investigated a limited number of studies which hindered the interpretation of bias.

## CHAPTER 6. DISCUSSION

### Summary of Findings

This research study had two purposes: to explore the impact of IGP on older adults' self-esteem, life satisfaction, depressive symptoms, and psychological well-being, and to examine potential moderators of IGP. The results suggest that overall, IGP participation was associated with beneficial effects on targeted well-being outcomes for older adults. Exploration analyses of moderators revealed several potentials for future research. The following section presents a summary and interpretation of findings, implications as well as limitations of the current study and suggestions for future research.

The first aim was composed of one research question: how effective was IGP to older adults. A Hunter and Schmidt meta-analysis approach (2014) was conducted and the findings suggested that IGP was effective but the effect was variable, implying the possibility of moderators. The second aim consisted of four research questions. Various analyses including scatterplots, correlation, ANOVA, independent samples *t*-test, and sample-weighted regression were run. The aim was to investigate possible moderators through comparing IGP effectiveness regarding various IGP characteristics. Regarding the serving type, I found that serving each other and younger individuals serving older adults had higher effectiveness than programs where older adults were serving younger individuals. In terms of the ratio of older to younger participants and program length, findings did not provide evidence for an association between each variable and the program effectiveness. Younger participants' age was highly and positively correlated with IGP effectiveness. Further exploration showed, younger individuals' age was critical to the success of IGP. A good practice of IGP needs to be a serving type of serving each other or

younger serving older, not too lengthy, with program support, and incorporate relatively “older” younger participants (e.g., college-aged individuals).

### **Implications**

#### **IGP duration**

In terms of interval between two sessions and the effect size, I found that most studies had older and younger participants meet once a week. From the scatterplots, it appears that meeting once a week was more beneficial to older adults, while when younger and older met too frequently (more frequently than once a week), a program became less effective, which is illustrated in Figure 6.

For the length of intervention per session, even though the effect sizes widely varied, it was noticeable from the scatterplots that the most effective meeting times lie between 45 minutes and two hours and that meeting less than 30 minutes per session was very short and rather ineffective. In terms of program length, while the correlation between the IGP effect and the length of the program was not significant, the negative value signified that the longer the program, the less effective the IGP. In addition, from the scatterplots, it seemed that a program lasting within six months would allow for prime levels of effectiveness. It appears possible to have a short and yet effective IGP.

As IGP was a social exchange process, the principles of equity theory and social exchange theory such as fair distribution and satisfaction level could be applied to explain why longer IGP did not work well. For example, in the early stage of IGP, the rewards such as perception exceeded cost and with IGP advancing, the cost such as time and energy exceeded rewards. In addition, participants' expectation is likely to increase as IGP was going on, such

that even IGPs brought more benefits and participants' perception enhanced, as their expectation increased, their satisfaction level was harder to be met.

To sum up, for IGP duration, I found that 1) meeting too frequently, for example more than twice in a week was less effective than once a week, 2) any time between 45 minutes to two hours per session was the ideal programming length per session, and 3) lengthy programs did not necessarily produce the best results. Again, due to the limited number of effect sizes, determining the best program length (the weeks IGP lasted for) and intervention time per session requires more evidence.

### **Serving type and ratio**

Based on the mean effect sizes from the three groups, it can be inferred that IGP of older adults serving younger individuals was less effective to older adults' psychological well-being. While both younger serving older and older and younger serving each other had higher effect sizes, serving each other appears to be the best serving type for older adults. This is consistent with our hypothesis that, based on equity theory and social exchange theory, serving each other is the most stable and beneficial serving type for both younger and older in the long run. It is not surprising to see that younger serving older had higher effect size for older adults since this serving type could maximize older adults' benefits.

There were incomplete data in terms of the ratio of older to younger individuals, as only nine of the 15 studies provided information that could be transformed into such ratios. One potential reason for so much omission could be that researchers paid inadequate attention to this information when they designed the study. For articles that provided ratio information, some researchers assigned all participants into one large group and some studies assigned participants into small groups. A rationale for group size was not presented by these authors. Based on equity



theory and social exchange theory, while a certain ratio of older to younger participants would represent an optimal ratio that would create the greatest benefit, the current study failed to find evidence to support that the ratio of older to younger was important to IGP success, and more research is needed to explore that ratio.

Based on the ratio of older to younger individuals in Table 2, it was evident that the more effective IGP has a ratio of older to younger individuals near one. By examining the two most effective IGPs to older adults (the 12<sup>th</sup> article and the 4<sup>th</sup> article), I found that the most effective group was with about 10 older adults and 10 younger individuals, but I suggest that future researchers report the number of older adults and younger individuals in a small group in IGP so that further conclusions may be drawn regarding the ratio of older to younger participants. A summary of the ratio of older to younger is outlined in Table 2 & 8 and Figure 6.

### **Older and younger participant ages**

As indicated from the correlation table, ages of the younger individuals were significantly and positively related to older adults' psychological well-being. A closer examination of Table 5 (the correlation table) showed that younger individuals' age was not significantly related to program support, IGP duration, or other IGP characteristics, indicating that the age of the younger individual was an important independent factor in determining the effect of IGP with respect to older adults' psychological well-being. More specifically, the scatterplots showed IGP to be less effective for preschool students while it was more effective for college-aged students.

Based on older adults' age and the effect size, it is noticeable that current IGPs work best for 70-80-year-old older adults, while it is less effective for 80-90-year-old adults on their psychological wellbeing. From Figure 6, it can be seen that only the group of 70-80-year-old interacted with college students, while those in the 60-70 age group had interaction only with

school-aged children and the 80-90 age group had interaction with children from preschool through high school. It is possible that the reason why the age group 70-80 had the best results is because of the younger individuals with whom each specific age group interacted.

The reason for these two phenomena (younger's age was positively related to the effect and 70-80 older adults had best IGP effect) could be that the older the younger individuals, the more interaction was ignited and equal exchange was made more accessible between younger individuals and older adults. In addition, if the older adults were relatively older (e.g., 80-90) than their peers, it is likely that their psychological well-being status is relatively lower, thus it is harder to yield much significant benefit for older adults through the interaction.

When connecting the above two phenomena (younger and older adults' ages and the combination of older and younger individuals) with Erikson's theory of psychological development, it appears reasonable that 70-80 year old groups had the highest IGP effect because they interacted with mainly secondary school students and college students, who have the need of identity formation (Erikson, 1963) and older adults are in need of generativity (Erikson, 1963) which contributes to the psychosocial well-being (Knight, Skouteris, Townsend & Hooley, 2014). The conclusion that the combination of oldest-older adults and the youngest-younger individuals had the least IGP effect also seemed reasonable. Per Biggs and Knox (2014), older adults expressed that younger individuals were loud while younger individuals stated that they learned to not play loudly. It appears that when oldest older adults and the youngest younger individuals interacted one another, their needs cannot be met through IGP, which resulted in the low IGP effect. However, due to the limited combination of age groups, an exploration of more age group combinations is needed before making determining conclusions.

### **Program support and control group**

The independent samples t-test (Table 6) and regression analysis (Table 7) showed even though the result was not significant, major difference in effect size between studies with and without program support existed. It indicated that program support was a critical factor in determining the effectiveness of IGP to older adults, which is consistent with previous research that emphasized continued support to participants (Dupuis, Wiersma, & Loiselle, 2012).

In terms of the difference between studies with and without control groups, the results (Table 6) showed studies with control groups had lower effect size compared with studies without control groups. There are two explanations for this unexpected phenomenon. First, as indicated by multiple authors, the control groups were intended to be randomly distributed but it was not random, so it is possible that participants in control groups were originally different from participants in experimental groups, which then caused the difference in our analysis. The second explanation was that IGP was not as effective as we expected. Since a design with a control group would be more reliable compared with designs without control groups, IGP with control groups may be more accurate and the analysis result maybe more reliable. Thus, standardized control groups are needed in the design of future IGP to more accurately assess the effectiveness of IGP.

Above all, IGP was found to be effective for older adults ( $d = 0.37$ ,  $SD=0.50$ ). IGP participation was associated with fewer depressive symptoms and increased self-esteem, life satisfaction, and psychological well-being among older adults. In terms of IGP characteristics, I concluded that the best IGP curricula for older adults should focus on younger individual age and program characteristics. In the current study, a list of younger individuals' effect sizes, program characteristics, and participant characteristics was outlined in Table 8. In view of the limited

number of studies, more research is needed to draw a generalized conclusion about the best IGP curricula for both younger and older participants.

### **Logic model**

A summary of research questions and interpretations is presented in Table 10. These findings are important for the future replication of IGPs. As confronted with the various characteristics of IGPs, it is hard to duplicate a successful IGP at other locations. For example, even though IGPs exist in several countries, I was not able to find one IGP article published in mainland of China, the country with the largest number of older adults in the world. Identifying a clear IGP model can help with duplication and development of IGPs around the world.

To identify the best IGP curricula for both younger and older participants, a logic model of IGP (Figure 7) was presented. “Logic Models are helpful to articulate their understanding of the current situation, the changes they hope to bring about, the activities planned to contribute toward this change, the resources needed to put into the effort, assumptions they are making, and external factors that could influence the results” (University of Wisconsin-Extension, 2016). The logic model provides the resources needed to promote the development and evaluation of IGP. Availability of IGP curriculum and best practices are needed. Currently available is the Bridges Program Curricula Suite created by Andrea J. Fonte Weaver, however, this curriculum does not appear to be empirically tested. IGP curricula that are evidence-based, free, and ready to be applied nationally and throughout the world, are needed. In addition, researchers can follow the evidence-based non-IGP programs such as Strengthening Families Programs to better develop IGP curriculum.

### **Meta-analysis bias**

In this study, the graphs of four funnel plots were skewed and asymmetrical, indicating the messiness and the presence of bias. It seems that these funnel plots include both published and unpublished, small samples and large sample sizes, and positive and negative studies. However, a larger sample size is needed to accurately determine the presence of bias in the future.

### **Limitations of the Current Study**

The following section notes the limitations of the current meta-analysis. In terms of screening articles, if the title contained “case study”, the article was excluded. This is a limitation because it is possible that a case study also provided pretest and posttest data sufficient to obtain the effect size, such as in a study involving three older adults and three younger individuals (Talbot, 2000). Another limitation is that when the title stated *younger individuals pretest and posttest attitude changes in IGP without mentioning older adults*, the article was excluded. It is possible that when the outcome for older adults was not significant, the author will not mention older adults in the title while still providing data regarding older participants.

Next, while I searched the list of the combined keywords in both the abstract and the title, I did not search for them as keywords in the paper. There may have been papers that only mentioned “intergenerational program” as paper keywords, but did not use these words again in the paper. For example, while a study investigated the impact of an altruistic activity on older adults’ life satisfaction was a typical IGP study (Yuen, 2003), the author only mentioned “intergenerational program” as the keywords of the paper but not in the text, so future researchers may also want to search the list of keywords as keywords of the studies.

There is no standardized coding manual, so our coding may not be the most-recognized type of coding. For example, I coded the serving type based on the activity content, and if the activity is one in which older adults served as mentors of younger individuals, I coded the serving type as “older serving younger”. The younger individual’s age was coded based on their age period or the age range. In terms of program support, only if the facilitator was encouraging younger and older participants to have more interaction was it considered program support. One example of coding limitation would be that I coded activity type as either personal related activity or social activity, which resulted in only one personal related activity out of the 15 studies, and no analysis was therefore performed.

In addition, although our focus was on healthy older adults, our criteria of “healthy older adults” was too broad. The current study ended up including older adults with slight depressive symptoms with a score of no more than 18 out of 30 on Geriatric Depression Scale (Hernandez, 2008) and slightly cognitively-impaired older adults whose condition did not affect their interaction with others were also included in the current study.

Finally, another limitation was related to the analysis. The approach to handling missing data was pairwise deletion, leading to fewer effect sizes included in the correlation and regression. Also, due to the limited number of articles in each targeted outcome, I integrated life satisfaction, self-esteem, and depressive symptoms as the construct of psychological well-being. Even though previous literature suggests that life satisfaction, self-esteem, and depressive symptoms were highly correlated, it is essential to point out the three aspects do not represent all potential aspects of psychological well-being, and the three constructs are not the same in terms of the sensitivity to change confronting the IGP intervention.

### **Limitations of Included Studies**

The finding of the current research is significant in that it showed IGP increased older adults' life satisfaction, self-esteem, psychological well-being, and decreased depressive symptoms and suggested the possibility of moderators. However, this study also unveiled multiple limitations in conducting IGP meta-analysis, this section provides limitations I encountered and recommendations for future IGP researchers.

First, even though previous researchers had done a great deal of work in this field, a close examination of previous research showed that the definition of IGP was broad and unclear. For example, younger and older individuals together without interaction was considered as an IGP or an IGP study did not consider itself to be an IGP but rather a different intervention such as altruistic activity in older adults (Yuen, 2003), so future studies examining IGP in a more scientific, well-operationalized, and clearly-defined approach is needed.

Next, a final number of 15 articles selected from more than ten thousand articles indicated that most articles had limitations that prevented them from being included in the current meta-analysis. For example, there were many articles that did not meet the inclusion criteria such as targeted outcomes and limited spanned periods. In addition, there were more papers examining older adults with dementia rather than healthy older adults, and among those that studied healthy older adults, many examined shared-site IGP and familial IGP. As suggested by Knight et al., (2014), 584 out of 1,337 articles were familial IGP, 391 out of the 487 did not report outcome measures, and 82 out of the 96 did not mention reciprocal benefits. In addition, when the sample size was too small, the Hunter and Schmidt meta-analysis program cannot produce valid outcome information. Further, there were multiple limitations related to inconsistent reporting

and incomplete reporting of the findings and Table 9 provides a detailed summary of difficulties in performing a meta-analysis for IGPs.

In addition, the keywords of the current study [(“Intergeneration\*” OR “trans generation\*” OR “cross age” OR “cross generation\*” OR program\*), (“old\* adult” OR senior\* OR elder\*), and (depress\* OR “self-esteem” OR “life satisfaction” OR “quality of life”)] yielded articles that focused on programs of senior supporting senior and younger teaching younger. One example would be a study evaluating the senior companion program (senior supporting senior) instead of an IGP (Butler, 2006). Efforts are needed on developing the current keywords in such a way that only studies including both younger and older participants are retrieved.

Another limitation is that most experimental studies included in the current meta-analysis were not of random distribution. As indicated by several authors, there was an initial goal of random assignment of the participants but the actual design was not randomly assigned. The way of random assignment is able to balance the participant characteristics in different groups; however, this is not true in the studies included in the current analysis. In addition, participants in most IGP studies were recruited as convenient samples rather than random samples so the interpretation of conclusions cannot be generalized.

Further, since most IGP studies used clustered samples meaning the studies recruit participants from a single nursing home or a single community and all eligible participants in this cluster were included in the study, severe issues arise. For example, this clustered sample is only representative of this chosen nursing home or community (Himelein, Eckman, & Murray, 2013). One issue is that the effect sizes tend to be larger than the real effect size because participants’ experience in the control group and intervention group may influence the mental effect which contributes to the difference of the effect size between control group and the intervention group.



Resentful demoralization is an issue related to the clustered samples. “If subjects learn that the treatment level to which they have been assigned received less desirable goods or services, they may experience feelings of resentment and demoralization” (Kirk, 2014, p.22). For example, in one study, older adults in the control group were told that there were not enough children to interact with. “Their response may be to perform at an abnormally low level, thereby increasing the magnitude of the difference between their performance and that of subjects assigned to the desirable treatment level” (Kirk, 2014, p.22). These two factors both cause systematic bias when evaluating IGP effectiveness.

Another limitation is that only participants who remained in IGP were included in the analyses, which will lead to systematic bias. Thus, the analysis only included participants who keep benefiting from IGP or only included successful IGPs. In another word, participants remained in IGP might because they constantly experienced benefits from IGP or simply because the IGP they participated in was working well. In addition, when participants were diagnosed with certain emotional or cognitive disorders or changed between control group and intervention group as suggested by Murayama et al. (2015), only highly functioning participants were left in the program such that the study will lead to more selective results.

### **Recommendations for Future Research**

First, to modify the searching terms so that only articles containing both younger and older adults are retrieved. One way to modify the keywords would be combining “younger, youth, kids, and child\*” in our list of keywords. Next, consistent with Jarrott’s (2011) suggestion that IGP need improvements in the following fields: standardized measures and rigorous application of qualitative and quantitative analyses, I have summarized recommendations for future researchers in conducting experimental or quasi-experimental IGP research as follows.

The first thing to consider in conducting experimental or quasi-experimental IGP research is to design a more scientific program, which is consistent with the suggestion of Jarrott (2011) that more rigorous researches are needed. One way of conducting a more scientific research is to include a control group along with the experimental group, meaning that participants in the control group have no intergenerational contact during the experiment. A control group could help researchers to investigate the “effect of various factors one by one and gain a clear picture of the factors that actually contribute to the phenomenon” (Shaughnessy & Zechmeister, 1994, p.9). Another approach is to employ random assignment and the random sample to improve both internal and external validity (Shaughnessy & Zechmeister, 1994).

Next, a more scientific intervention is also needed for the experimental group, or else it will be difficult to understand the true experimental effect. For example, a group of older adults in the experimental group met with younger individuals from kindergarten, preschool, primary school and middle schools and were simultaneously engaged in various activities with other older adults, so it was impossible to know which intervention is causing the effect since it could have been infants with older peers, primary-school students with older adults, or other combinations of younger and older adults. Other aspects studies need to report more scientific studies include the training to gold standard in conducting the experiments and widely recognized IGP curriculum, and assess and report fidelity to treatment/program/intervention.

In addition, it is necessary for future researchers to adopt a more scientific approach to reporting IGP findings. For example, if a researcher has reported only the posttest mean and standard deviation of the experimental group, the effectiveness value of the program was then inaccessible. Furthermore, more information about the sample characteristics should be reported in future studies. Currently, in this meta-analysis, five out of 15 articles provided an age range

without the mean age of the older participants. The average age was estimated by taking the mean of the age range to examine the correlation between older participants' age and the effect size, but knowing the actual mean age would be preferred.

In terms of reporting study findings in a more scientific way, I summarized the suggestions from APA publication and communications board that provide the reporting standards for psychology studies (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008) as follows. First, to provide more detailed, scientifically-based, and evidence-based information when reporting studies. Second, to comply to the principles: “consort” (random assignment) and “transparent” (transparent reporting of scientific studies). Specifically, I recommend future researchers strictly follow the journal article reporting standards and the meta-analysis reporting standards provided by APA Publication and Communications Board (2008).

Regarding scientific reporting of research findings, one important aspect is reporting attrition in pre- and post-experiments. Through this meta-analysis, it was noticed that many studies failed to compare the characteristics of participants who were included and excluded from the final analysis, which could bring in bias, especially if high attrition rates were observed. As indicated from the included IGP studies, shorter IGP had relatively low attrition rate as evidenced that most shorter IGP included had an attrition rate of less than 10 percent, however, those longer IGP had high attrition rate. For example, a program lasting for two years had an attrition rate of 41% (Murayama et al., 2015) and a program lasting for 7 years had an attrition rate of 54% (Sakurai et al., 2016). To provide more information of participants excluded from the analysis and identify potential attrition bias, it is suggested to present baseline characteristics

of those included in the analysis and those who are excluded from the analysis separately (Dumville, Torgerson, & Hewitt, 2006).

While the core of the current meta-analysis is the coding of the articles, no existent coding manual is available. I suggest that future researchers create a coding manual to record how the program characteristics and participant characteristics were coded. A widely-recognized IGP coding manual would help IGP development by ensuring that all IGP designs will have important IGP characteristics to support consistent and scientific development of IGP. In addition, explicit and specific coding instructions will help the definition of IGP become more accurate, less broad, and clearer. Figure 7 outlines what is needed in intergenerational programming including inputs (participants, setting, theory, goals), outputs (program support, program, program length, the ratio), and outcomes (short-term outcomes, medium-term, long-term outcomes).

Faced with the various outcome measures, I suggest that future researchers compare different measurement scales and present the correlations of outcome variables, making it possible to merge highly-correlated variables such as self-esteem and general self-efficacy. Of note, certain measurement scales such as depressive symptoms are less likely to show change compared with measurement scales like life satisfaction. Therefore, future researchers may consider combining measurement scales of similar traits, for example, combining outcome variables of long-term health and short-term health separately, and combining variables of health behaviors and health outcomes separately. Furthermore, future researchers should be consistent with respect to the measurement scales used in pretest and posttest.

In terms of the limited combination types of younger and older adults' ages, I suggest that researchers recruit a large number of participants and both younger and older participants need to

cover a various age range. For example, younger individuals contain both primary school students and college students and older adults consist of younger old, old, and oldest old. Such that, older adults in each age group have interaction with younger individuals within each age group, which allows for the in-depth examination of various age combinations of younger individuals and older adults.

Another recommendation for future researchers is related to the measurement reliability. It is important that authors employ measurement with high reliability and report the reliability of a measurement from the current sample. As for Hunter and Schmidt's meta-analysis software (2004), incomplete knowledge of reliability would result in defective calculation of the effect size. In addition, studies not employing a reliable scale may ask questions that create low reliability which also influences the effect size.

IGP results in both direct and indirect outcomes for older adults. As indicated from our theoretical model, IGP first affects direct outcomes like aging attitudes and stereotypes, then influences indirect outcomes such as health-related outcomes. While direct effects are important, more research is needed to investigate the indirect effect of IGP, because the main purpose of IGP is not solely to increase interaction between younger and old, but more importantly to enhance older adults' health-related outcomes. Future researchers should continue shedding light on this area, specifically with respect to health-related outcome variables, which is consistent with Jarrott's (2011) recommendation that researchers need to employ measurements beyond perceptions and attitudes.

Finally, I suggest that future researchers investigate the effect of IGP on both younger and older adults simultaneously because IGP characteristics may be different when older adults benefit the most as a group while younger individuals benefit the most individually. It is critical

to identify effective IGP practice (valid IGP characteristics) because it is advisable for policy makers to focus on such valid practices and to assign more resources to sustain successful IGP while terminating or modifying ineffective IGP, because the current research demonstrated that, while IGP could be very effective, it could also be ineffective.

## **Conclusions**

The first purpose of this study was to investigate the effectiveness of IGP on older adults' psychological well-being as indicated by self-esteem, depressive symptoms, and life satisfaction. The second purpose was to explore the existence of potential moderators. Findings from the meta-analysis suggest that IGP improved older adults' self-esteem, life satisfaction, depressive symptoms, and psychological well-being and revealed the existence of moderators in all four outcomes. Exploratory analyses indicated that younger individuals' age and other IGP characteristics such as having program support were possible moderators. An important next step would be to incorporate more studies to identify potential moderators of IGP effectiveness and investigate the effectiveness of IGP for both older adults and younger individuals.

## REFERENCES

- Adams, J. S. (1963) Toward an understanding of inequity. *Journal of Abnormal and Social Psychology*, 67, 422-436
- Anderson, K., Anderson, L. E., & Glanze, W. D. (1998). *Mosby's Medical, Nursing, & Allied Health Dictionary*. CV, Mosby: Mosby-Year Book.
- APA publications and communications board working group on journal article reporting standards (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *The American Psychologist*, 63(9), 839-851.  
doi:10.1037/0003-066X.63.9.839
- Battle, J. (1978). Relationship between self-esteem and depression. *Psychological Reports*, 42(3), 745-746. doi:10.2466/pr0.1978.42.3.745
- Biggs, M. J. G., & Knox, K. S. (2014). Lessons learned from an intergenerational volunteer program: A case study of a shared-site model. *Journal of Intergenerational Relationships*, 12(1), 54-68. doi:10.1080/15350770.2014.869981
- Branden, N. (1971). *The psychology of self-esteem: A new concept of man's psychological nature*. Bantam Books.
- Breytspraak, L. M., Arnold, L., & Hogan, K. (2008). Dimensions of an intergenerational relationship between medical students and mentors-on-aging. *Journal of Intergenerational Relationships*, 6(2), 131-153. doi:10.1080/15350770801955040
- Brown, G. W., Andrews, B., Harris, T., Adler, Z., & Bridge, L. (1986). Social support, self-esteem and depression. *Psychological Medicine*, 16(4), 813-831.  
doi:10.1017/S0033291700011831



- Butler, S. S. (2006). Evaluating the senior companion program: A mixed-method approach. *Journal of Gerontological Social Work*, 47(1-2), 45-70. doi:10.1300/J083v47n01\_05
- Cardona, R. (2002). *Task oriented intergenerational program: The effects on self-esteem and self-efficacy in high school students and self-esteem and depression in older adults* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. EP14554)
- Carstensen, L. L., Isaacowitz, D. M., Charles, S. T. (1999). "Taking time seriously: A theory of socioemotional selectivity". *American Psychologist*, 54 (3): 165. PMID 10199217. doi:10.1037/0003-066X.54.3.165.
- Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, 4(1), 62-83. doi:10.1177/109442810141004
- Chung, J. C. (2009). An intergenerational reminiscence program for older adults with early dementia and youth volunteers: Values and challenges. *Scandinavian Journal of Caring Sciences*, 23(2), 259-264. doi:10.1111/j.1471-6712.2008.00615. x.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Earlbaum Associates.
- Crede, M. (2016, September 26). Personal interview.
- Cross Validated. (2013, April 13). Is it possible to find the combined standard deviation? Message posted to <https://stats.stackexchange.com/questions/55999/is-it-possible-to-find-the-combined-standard-deviation>

- DeMichelis, C., Ferrari, M., Rozin, T., & Stern, B. (2015). Teaching for wisdom in an intergenerational high-school-English class. *Educational Gerontology*, 41(8), 551-566. doi:10.1080/03601277.2014.994355
- Deutchman, D. E., Bruno, K. A., & Jarrott, S. E. (2003). Young at heart: Intergenerational activities involving persons with dementia. *Activities Directors' Quarterly for Alzheimer's & Other Dementia Patients*, 4(2), 27-35.
- DeVore, S., Winchell, B., & Rowe, J. M. (2016). Intergenerational programming for young children and older adults: An overview of needs, approaches, and outcomes in the United States. *Childhood Education*, 92(3), 216-225. doi:10.1080/00094056.2016.1180895
- Diener, E., & Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal of Personality and Social Psychology*, 68(4), 653. doi:10.1037/0022-3514.68.4.653
- Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71-75. doi:10.1207/s15327752jpa4901\_13
- Doll, G., & Bolender, B. (2010). Age to age: Resident outcomes from a kindergarten classroom in the nursing home. *Journal of Intergenerational Relationships*, 8(4), 327-337. doi:10.1080/15350770.2010.520614
- DuBois, D., Felner, R., Brand, S., Phillips, R., & Lease, A. (1996). Early adolescent self-esteem: A developmental-ecological framework and assessment strategy. *Journal of Research on Adolescence*, 6(4), 543-579.

- Dumville, J. C., Torgerson, D. J., & Hewitt, C. E. (2006). Research methods: reporting attrition in randomized controlled trials. *BMJ: British Medical Journal*, 332(7547), 969-971. doi: 10.1136/bmj.332.7547.969
- Dupuis, S. L., Wiersma, E., & Loiselle, L. (2012). Pathologizing behavior: Meanings of behaviors in dementia care. *Journal of Aging Studies*, 26(2), 162-173. doi:10.1016/j.jaging.2011.12.001
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629-634. doi:10.1136/bmj.315.7109.629
- Erikson, E. H. (1963). *Youth: Change and challenge*. New York: Basic Books.
- Family Caregiver Alliance (2016). *Caregiver's guide to understanding dementia behaviors*. Retrieved from <https://www.caregiver.org/caregivers-guide-understanding-dementia-behaviors>
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Gaggioli, A., Morganti, L., Bonfiglio, S., Scaratti, C., Cipresso, P., Serino, S., & Riva, G. (2014). Intergenerational group reminiscence: A potentially effective intervention to enhance elderly psychosocial wellbeing and to improve children's perception of aging. *Educational Gerontology*, 40(7), 486-498. doi:10.1080/03601277.2013.844042
- Galbraith, B., Larkin, H., Moorhouse, A., & Oomen, T. (2015). Intergenerational programs for persons with dementia: A scoping review. *Journal of Gerontological Social Work*, 58(4), 357-378. doi:10.1080/01634372.2015.1008166

- Generations United (2007). *Fact sheet: The benefits of intergenerational programs*. Retrieved from <http://www.gu.org/LinkClick.aspx?fileticket=71wHEwUd0KA%3D&tabid=157&>
- Goyer, A. (2001). Intergenerational shared site and shared resource programs: Current models. (Available from Generations United, 122 C Street, NW, Suite 820, Washington, DC 20001-2109)
- Harris, P. B., & Caporella, C. A. (2014). An intergenerational choir formed to lessen Alzheimer's disease stigma in college students and decrease the social isolation of people with Alzheimer's disease and their family members: A pilot study. *American Journal of Alzheimer's Disease & Other Dementias*, 29(3), 270-281.  
doi:10.1177/1533317513517044
- Headey, B., Kelley, J., & Wearing, A. (1993). Dimensions of mental health: Life satisfaction, positive affect, anxiety and depression. *Social Indicators Research*, 29(1), 63-82.  
doi:10.1007/BF01136197
- Health-Related Quality of Life (HRQOL) (2016). *Centers for Disease Control and Prevention*. Retrieved from <https://www.cdc.gov/hrqol/concept.htm>
- Hernandez, C. R., & Gonzalez, M. Z. (2008). Effects of intergenerational interaction on aging. *Educational Gerontology*, 34(4), 292-305. doi:10.1080/03601270701883908
- Himelein, K., Eckman, S., & Murray, S. (2013). The use of random geographic cluster sampling to survey pastoralists. *World Bank Policy Research Working Paper No. 6589*. Retrieved from: <https://ssrn.com/abstract=2320861>
- HLWIKI International (2016). Scoping reviews. Retrieved from [http://hlwiki.slais.ubc.ca/index.php/Scoping\\_reviews](http://hlwiki.slais.ubc.ca/index.php/Scoping_reviews)

- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage Publications.
- Isaki, E., & Harmon, M. T. (2015). Children and adults reading interactively: The social benefits of an exploratory intergenerational program. *Communication Disorders Quarterly*, 36(2), 90-101. doi:10.1177/1525740114529154
- Jarrott, S. E. (2011). Where have we been and where are we going? Content analysis of evaluation research of intergenerational programs. *Journal of Intergenerational Relationships*, 9(1), 37-52.
- Jarrott, S. E., & Bruno, K. (2007). Shared site intergenerational programs: A case study. *Journal of Applied Gerontology*, 26(3), 239-257. doi:10.1177/073346480730022
- Jarrott, S. E., & Smith, C. L. (2010). The complement of research and theory in practice: Contact theory at work in nonfamilial intergenerational programs. *The Gerontologist*, 51(1), 112-121. doi:10.1093/geront/gnq058
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology*, 83(3), 693-710. doi:10.1037//0022-3514.83.3.693
- Kalisch, H. R., Coughlin, D. R., Ballard, S. M., & Lamson, A. (2013). Old age is a part of living: Student reflections on intergenerational service-learning. *Gerontology & Geriatrics Education*, 34(1), 99-113. doi:10.1080/02701960.2012.753440
- Kamei, T., Itoi, W., Kajii, F., Kawakami, C., Hasegawa, M., & Sugimoto, T. (2011). Six month outcomes of an innovative weekly intergenerational day program with older adults and

- school-aged children in a Japanese urban community. *Japan Journal of Nursing Science*, 8(1), 95-107. doi:10.1111/j.1742-7924.2010.00164.x
- Keil, L. J., & McClintock, C.G., (1983). A developmental perspective on distributive justice. In D. M. Messick & K. S. Cook (Eds.), *Equity theory: Psychological and sociological perspectives* (pp. 13-47). New York, US: Praeger Publishers.
- Kirk, R. E. (2014). *Experimental design: Procedures for the behavioral sciences*. Thousand Oaks, CA: SAGE Publications, Inc.
- Knight, T., Skouteris, H., Townsend, M., & Hooley, M. (2014). The act of giving: A systematic review of nonfamilial intergenerational interaction. *Journal of Intergenerational Relationships*, 12(3), 257-278. doi:10.1080/15350770.2014.929913
- Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukophadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being. *American Psychologist*, 53(9), 1017-1031. doi:10.1037/0003-066X.53.9.1017
- Krogstad, J. M. & Cohn, D. (March, 14<sup>th</sup>, 2014). *U.S. Census looking at big changes in how it asks about race and ethnicity*. Retrieved from <http://www.pewresearch.org/fact-tank/2014/03/14/u-s-census-looking-at-big-changes-in-how-it-asks-about-race-and-ethnicity/>
- Lee, M. M., Camp, C. J., & Malone, M. L. (2007). Effects of intergenerational Montessori-based activities programming on engagement of nursing home residents with dementia. *Clinical Interventions in Aging*, 2(3), 477-483.
- Low, L. F., Russell, F., McDonald, T., & Kauffman, A. (2015). Grandfriends, an intergenerational program for nursing-home residents and preschoolers: A randomized trial. *Journal of Intergenerational Relationships*, 13(3), 227-240.

- McCrea, J. M., Smith, T. B. (1997). Types and models of intergenerational programs. In S. Newman, C. R. Ward, T. B. Smith, J. O. Wilson, J. M. McCrea, G. Calhoun, & E. Kingson (Eds.), *Intergenerational programs: Past, present, and future* (pp. 117-125). Washington, DC: Taylor & Francis.
- Mendis, K. P. (1993). *The effects of participation in an intergenerational program on the psychological well-being of the elderly*. ETD Collection for Fordham University. AAI9313767. <http://fordham.bepress.com/dissertations/AAI9313767>
- Morita, K., & Kobayashi, M. (2013). Interactive programs with preschool children bring smiles and conversation to older adults: Time-sampling study. *BMC Geriatrics*, 13(1), 111-118. doi:10.1186/1471-2318-13-111
- Murayama, Y., Ohba, H., Yasunaga, M., Nonaka, K., Takeuchi, R., Nishi, M., ... & Fujiwara, Y. (2015). The effect of intergenerational programs on the mental health of elderly adults. *Aging & Mental Health*, 19(4), 306-314. doi:10.1080/13607863.2014.933309
- Neugarten, B. L., Havighurst, R. J., & Tobin, S. S. (1961). The measurement of life satisfaction. *Journal of Gerontology*, 16, 134 -143
- Newman, S. (1997). History and evolution of intergenerational programs. In S. Newman, C. R. Ward, T. B. Smith, J. O. Wilson, J. M. McCrea, G. Calhoun, & E. Kingson (Eds.), *Intergenerational programs: Past, present, and future* (pp. 117-125). Washington, DC: Taylor & Francis.
- Newman, S., & Hatton-Yeo, A. (2008). Intergenerational learning and the contributions of older people. *Ageing Horizons*, 8(10), 31-39.

- Newman, S., Ward, C. R., Smith, T. B., Wilson, J.O., McCrea, J.M., Calhoun, G., & Kingson, E., (1997). *Intergenerational programs: Past, present, and future*. Washington, DC: Taylor & Francis.
- Pavot, W., Diener, E., & Suh, E. (1998). The temporal satisfaction with life scale. *Journal of Personality Assessment*, 70(2), 340-354. doi:10.1207/s15327752jpa7002\_11
- Posada, M. M. (2006). *Effects of an intergenerational activity on nursing home residents: A children serving elders model* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3223917)
- Rich, P., Myrick, R., & Campbell, C. (1983). Changing children's perceptions of the elderly. *Educational Gerontology*, 9(5-6), 483-491. doi:10.1080/0380127830090512
- Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.
- Ruggiano, N. (2012). Intergenerational shared sites: An examination of socio-physical environments and older adults' behavior. *Research on Aging*, 34(1), 34-55. doi:10.1177/0164027511414916
- Sakurai, R., Yasunaga, M., Murayama, Y., Ohba, H., Nonaka, K., Suzuki, H., ... & Rebok, G. W. (2016). Long-term effects of an intergenerational program on functional capacity in older adults: Results from a seven-year follow-up of the REPRINTS study. *Archives of Gerontology and Geriatrics*, 64, 13-20. doi: 10.1016/j.archger.2015.12.005
- Salamon, M. J., & Conte, V. A. (1984). *Salamon-Conte Life Satisfaction in the Elderly Scale (LSES)*. Odessa, FL: Psychological Assessment Resources Inc.
- Salari, S. M. (2002). Intergenerational partnerships in adult day centers: Importance of age-appropriate environments and behaviors. *The Gerontologist*, 42(3), 321-333. doi:10.1093/geront/42.3.321



Shaughnessy, J.J., Zechmeister, E. B., (1994). *Research methods in psychology*. New York, US: McGraw-Hill.

Sherer, M., Maddox, J.E., Mercadante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R.W. (1982). The Self-Efficacy Scale: Construction and Validation. *Psychological Reports*, 51(2), 663-671. doi:10.2466/pr0.1982.51.2.663

Skropeta, C. M., Colvin, A., & Sladen, S. (2014). An evaluative study of the benefits of participating in intergenerational playgroups in aged care for older people. *BMC Geriatrics*, 14(1), 109. doi:10.1186/1471-2318-14-109

Snider, J. G., & Osgood, C. E. (1969). *Semantic differential technique: A sourcebook*. Chicago, IL: Aldine.

Stevens-Ratchford, R. G. (1993). The effect of life review reminiscence activities on depression and self-esteem in older adults. *American Journal of Occupational Therapy*, 47(5), 413-420. doi:10.5014/ajot.47.5.413

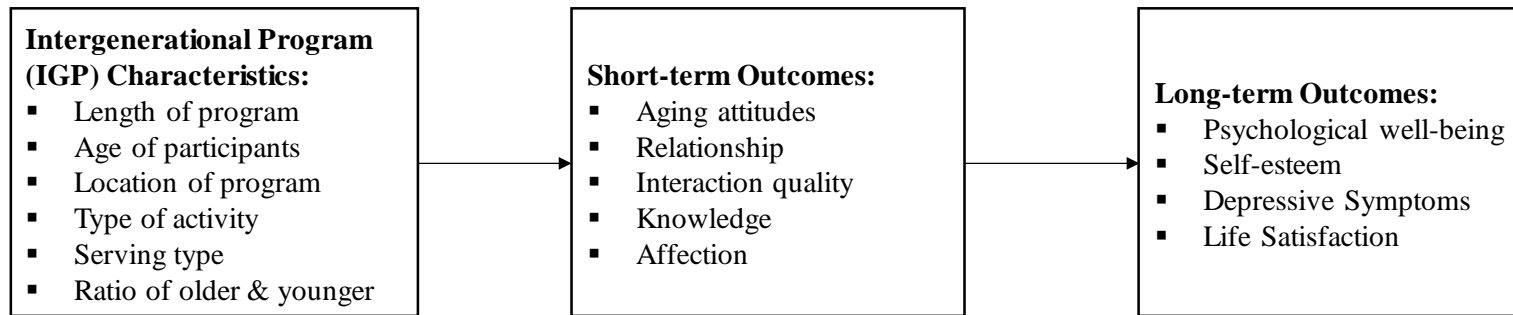
Sue, J. L. (2014). Do relationships matter? *An examination of a school-based intergenerational mentoring program* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. NS27992)

Swami, V., Chamorro-Premuzic, T., Sinniah, D., Maniam, T., Kannan, K., Stanistreet, D., & Furnham, A. (2007). General health mediates the relationship between loneliness, life satisfaction and depression. *Social Psychiatry and Psychiatric Epidemiology*, 42(2), 161-166. doi:10.1007/s00127-006-0140-5

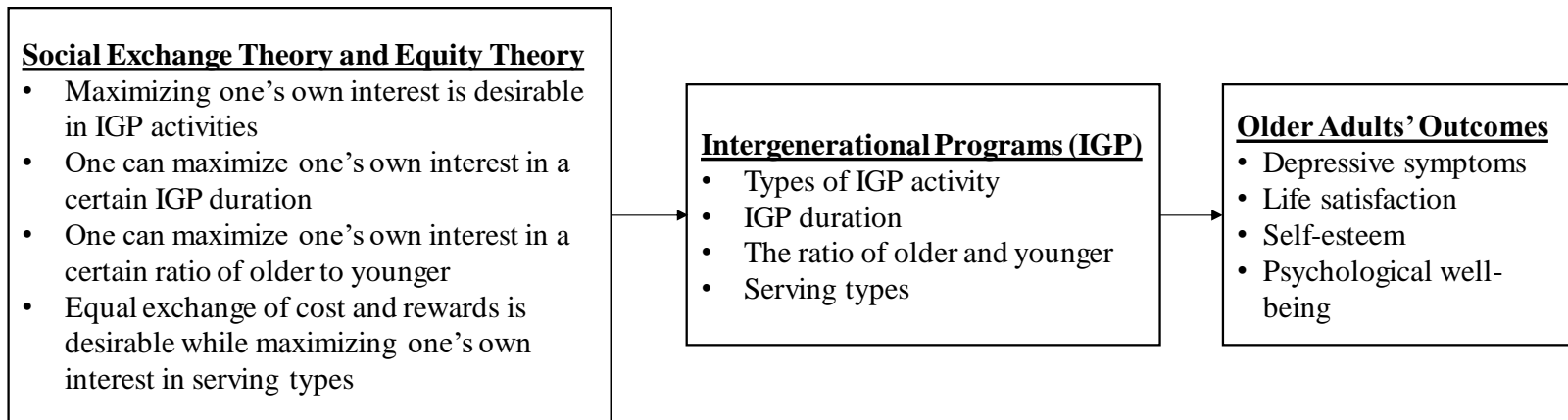
University of Wisconsin-Extension (2016). *Program development and evaluation: Logic Models*. Retrieved from <http://fyi.uwex.edu/programdevelopment/logic-models/>

- Ward, C. R. (1997). Evaluation of intergenerational programs. In S. Newman, C. R. Ward, T. B. Smith, J. O. Wilson, J. M. McCrea, G. Calhoun, & E. Kingson (Eds.), *Intergenerational programs: Past, present, and future* (pp. 117-125). Washington, DC: Taylor & Francis.
- West, M., & Hutchinson, S. (1992). Intergenerational geriatric remotivation. *Clinical Nursing Research*, 1(3), 221-235. doi:10.1177/105477389200100302
- White, J. M., Klein, D. M., Martin, T. F. (2015) *Family theories: An introduction (4<sup>th</sup> edition)*. Thousand Oaks, California: SAGE Publications, Inc.
- Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, 75(3), 315-321. doi:10.1037/0021-9010.75.3.315
- Williams, S., Renahan, E., Cramer, E., Lin, X., & Haralambous, B. (2012). 'All in a day's play' – An intergenerational playgroup in a residential aged care facility. *International Journal of Play*, 1(3), 250-263. doi:10.1080/21594937.2012.738870
- Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M.B., & Leirer, V.O. (1983). Development and validation of a Geriatric Depression Screening Scale: A preliminary report. *Journal of Psychiatric Research*, 17, 37-49.
- Yuen, H. K. (2003). Impact of an altruistic activity on life satisfaction in institutionalized elders: A pilot study. *Physical & Occupational Therapy in Geriatrics*, 20(3-4), 125-135. doi:10.1080/J148v20n03\_08

## FIGURES

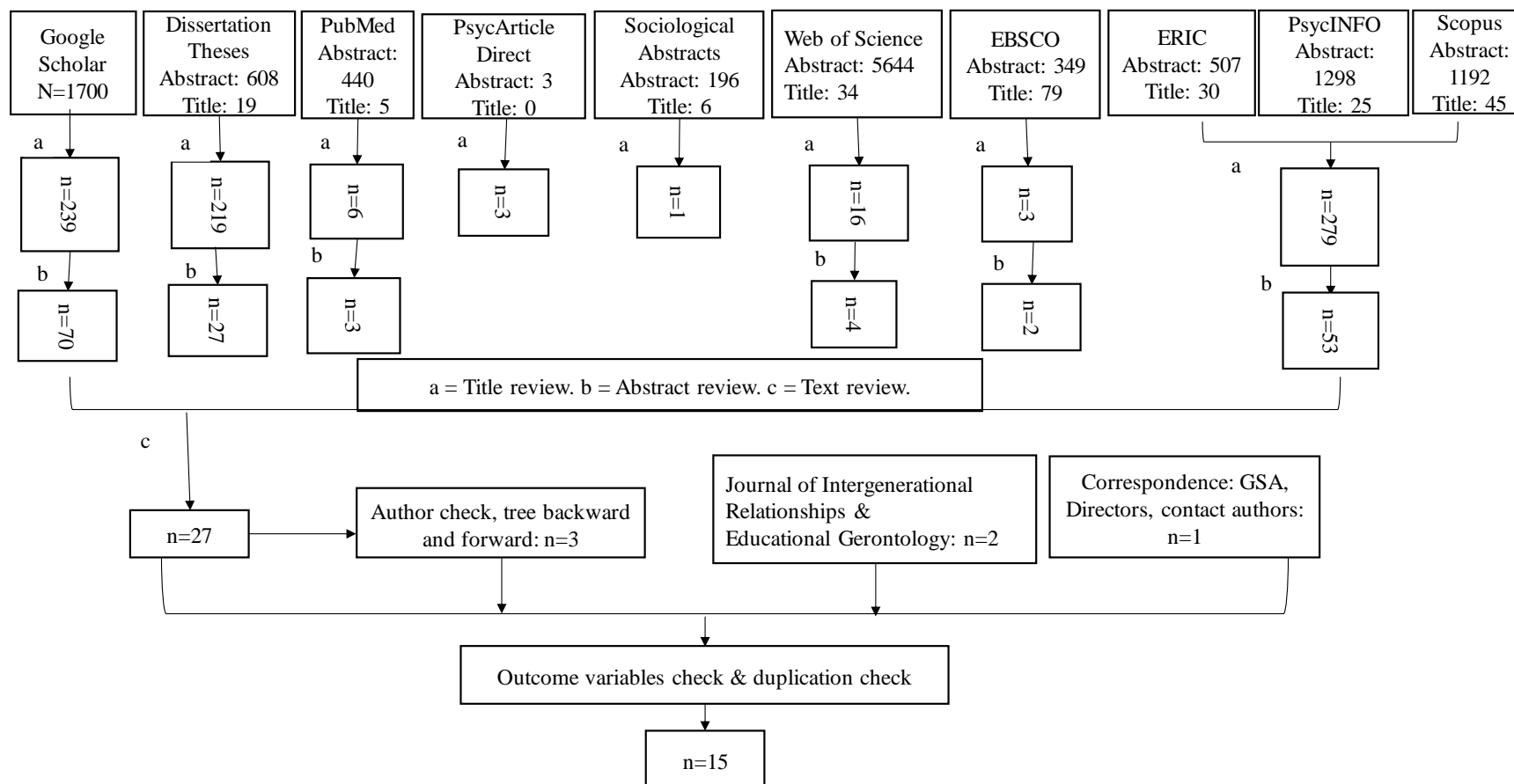


*Figure 1.* IGP characteristics and the effects for both younger and older participants. The effects include healthy behaviors (short-term effect of IGP for participants) and health and well-being outcomes (long-term effect of IGP for participants).



*Figure 2.* Connecting theories with program characteristics.

Based on social exchange theory and equity theory, an optimal level of IGP characteristics (activity type, IGP duration, the ratio, & serving type) determine the effectiveness of IGP for older adults through maximizing one's own benefits and obtaining equal exchange.



*Figure 3.* Article review process mainly included three resources: Ten databases, journals, and correspondence. The review process from the ten databases were illustrated through a, b, and c steps. Articles from journals and correspondence also went through the same process to the articles from ten databases: a, b, and c.

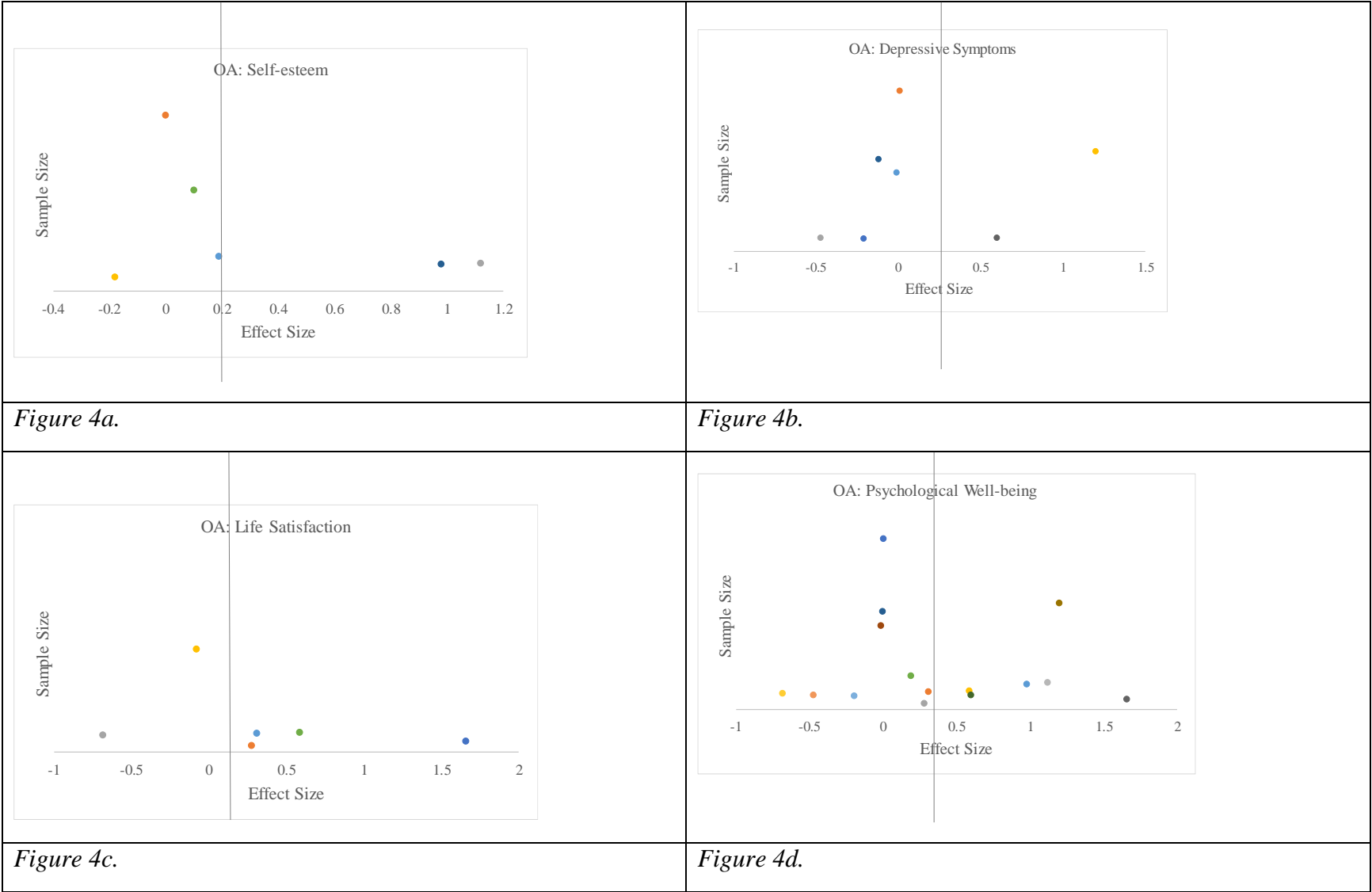
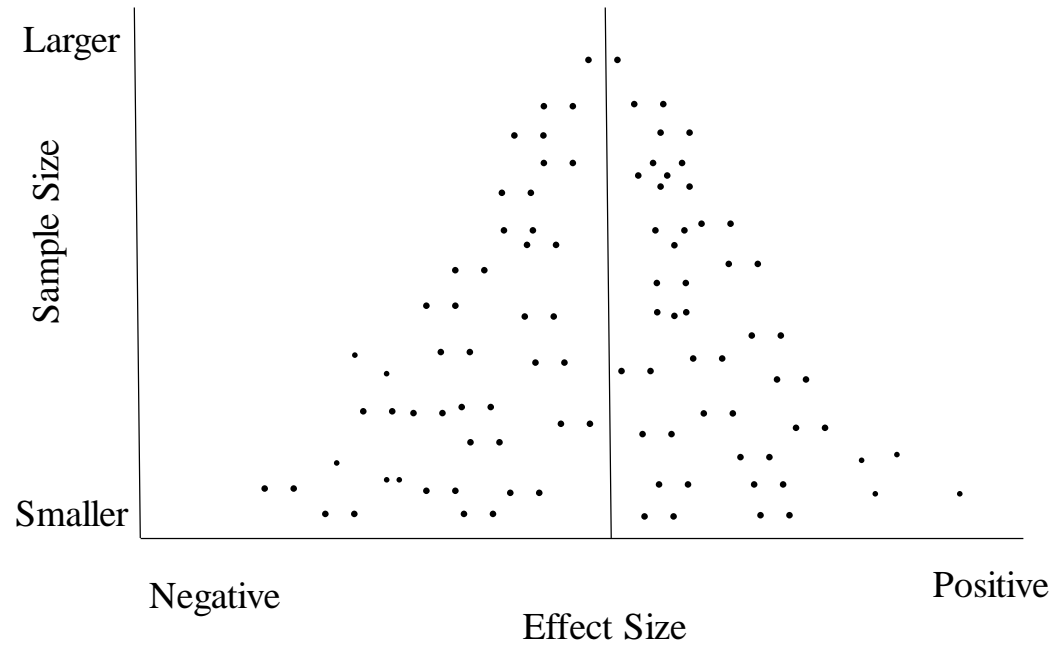
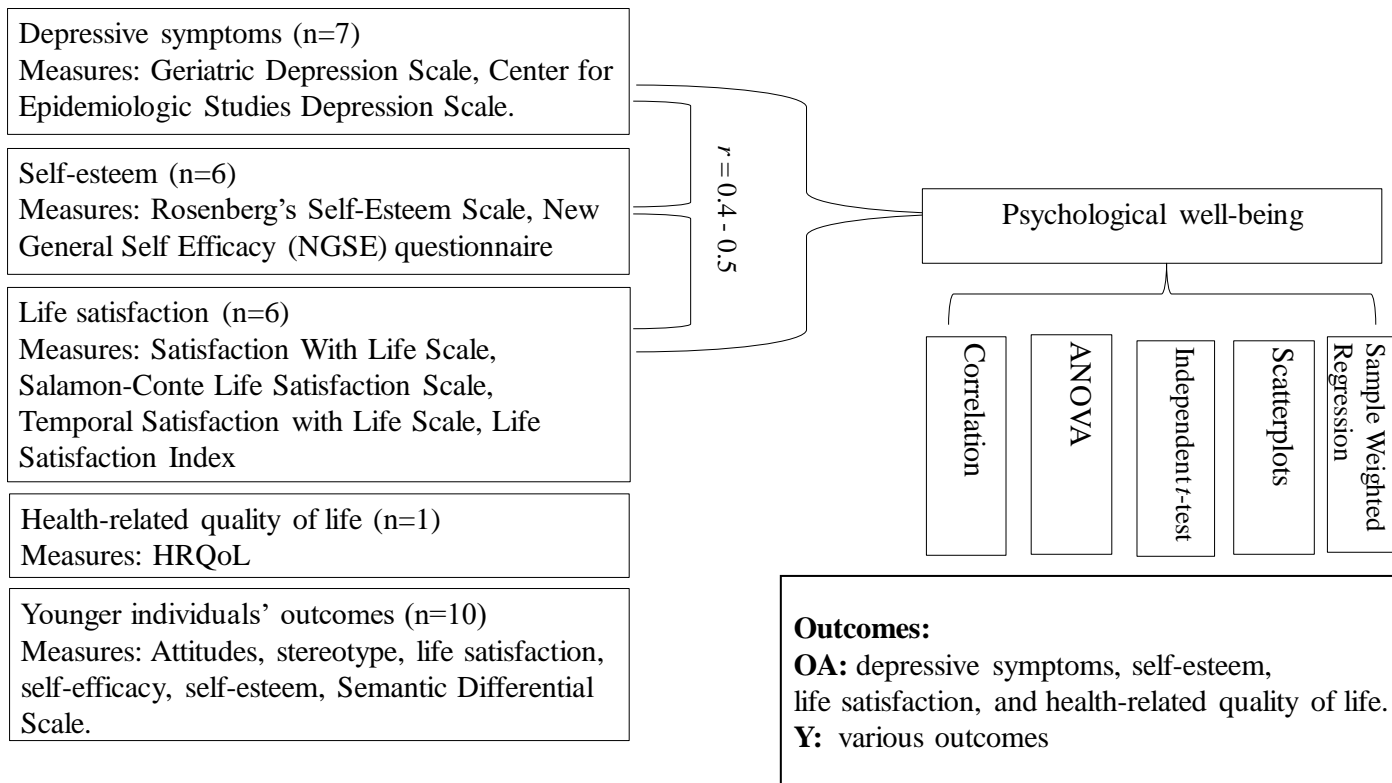


Figure 4. Funnel plots depicting number of participants and outcome effect sizes for each study: self-esteem (n=6), depression symptoms (n=8), life satisfaction (n=6) and psychological well-being (n=16). Each dot represents one unique study.



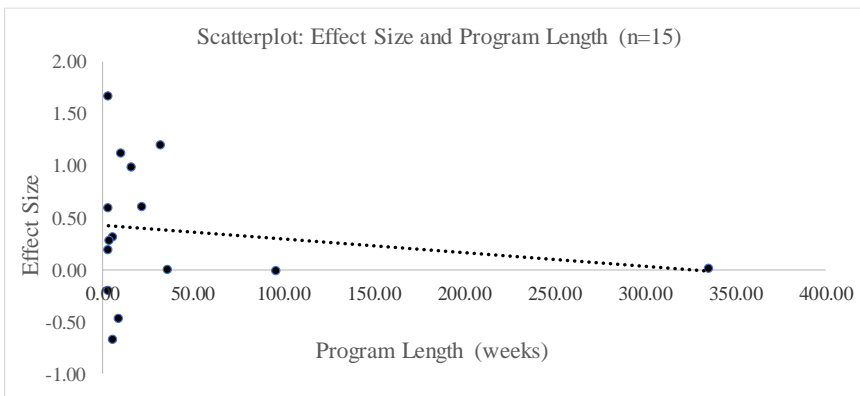
*Figure 4e.*

*Figure 4 (continued). Funnel plots depicting the study outcomes were compared to a standard symmetrical funnel plot in order to estimate degree of potential bias.*

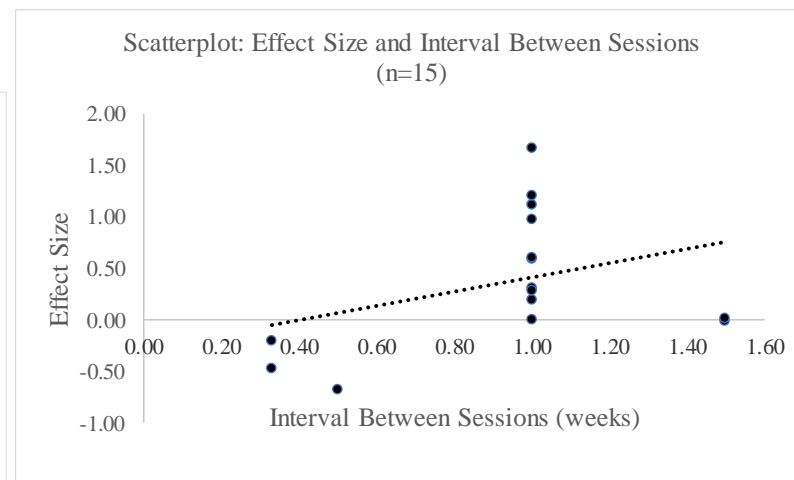


*Figure 5.* Combination of variables and specific analysis. All analyses were conducted using psychological well-being. Health-related quality of life was not included in the analyses as there was only one study. Younger individuals' outcomes were not involved in analyses due to the variety of outcomes.

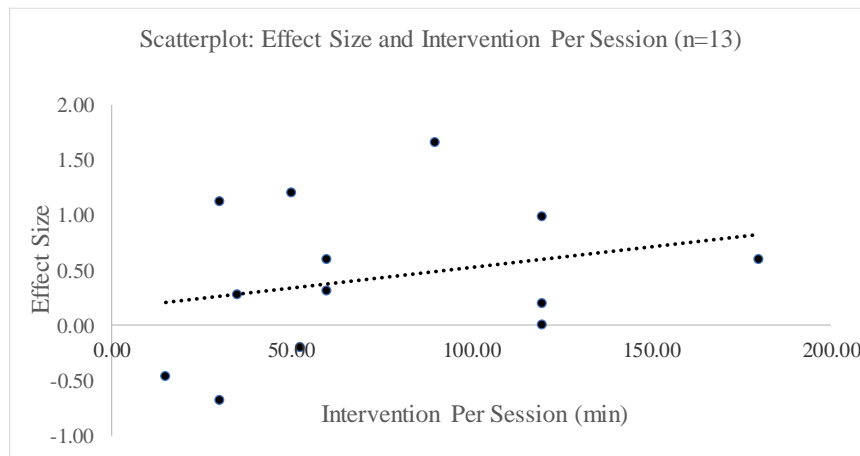




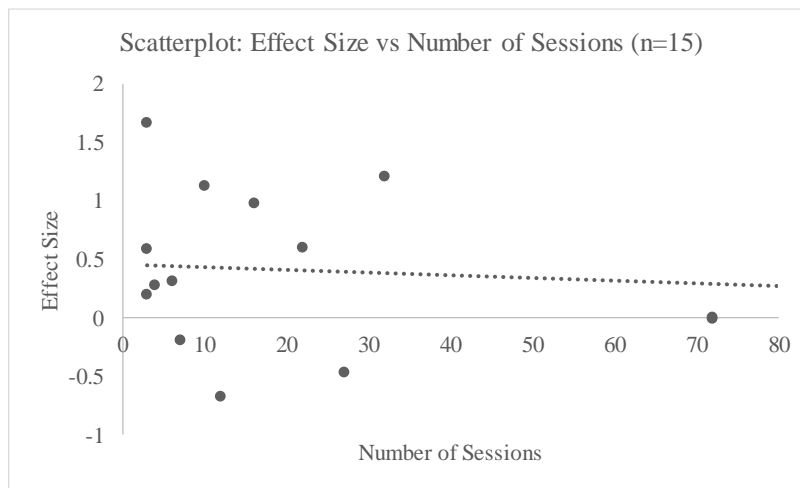
Panel A



Panel B

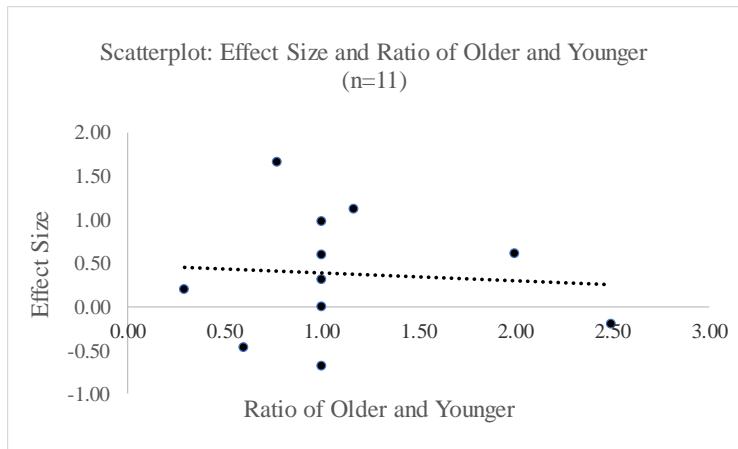


Panel C

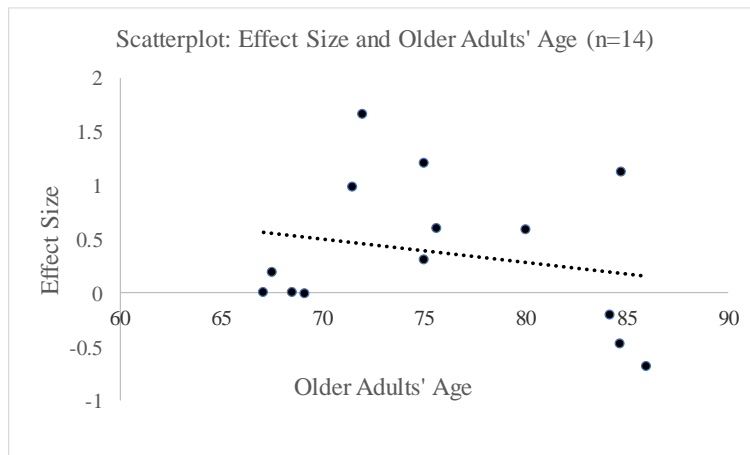


Panel D

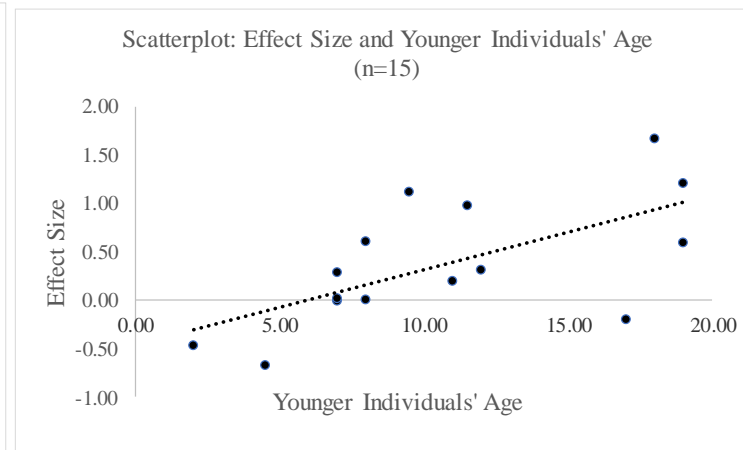
Figure 6. Association between effect sizes and program characteristics.



Panel E

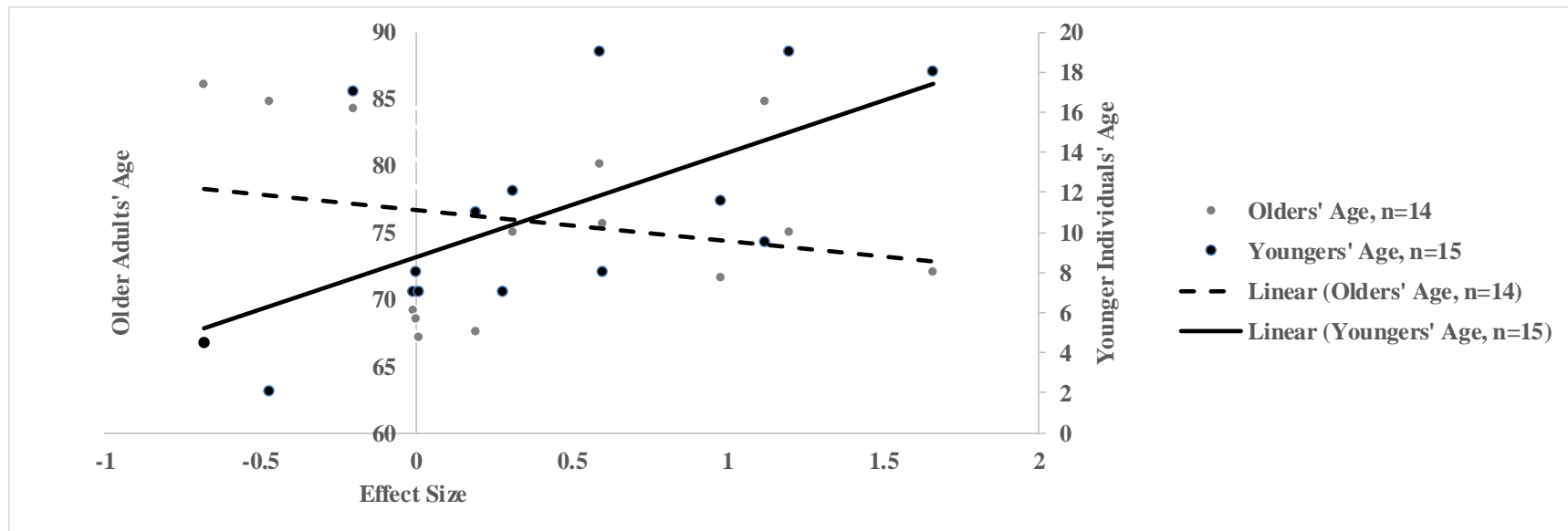


Panel F



Panel G

Figure 6 (continued).



Panel H

Figure 6 (continued).

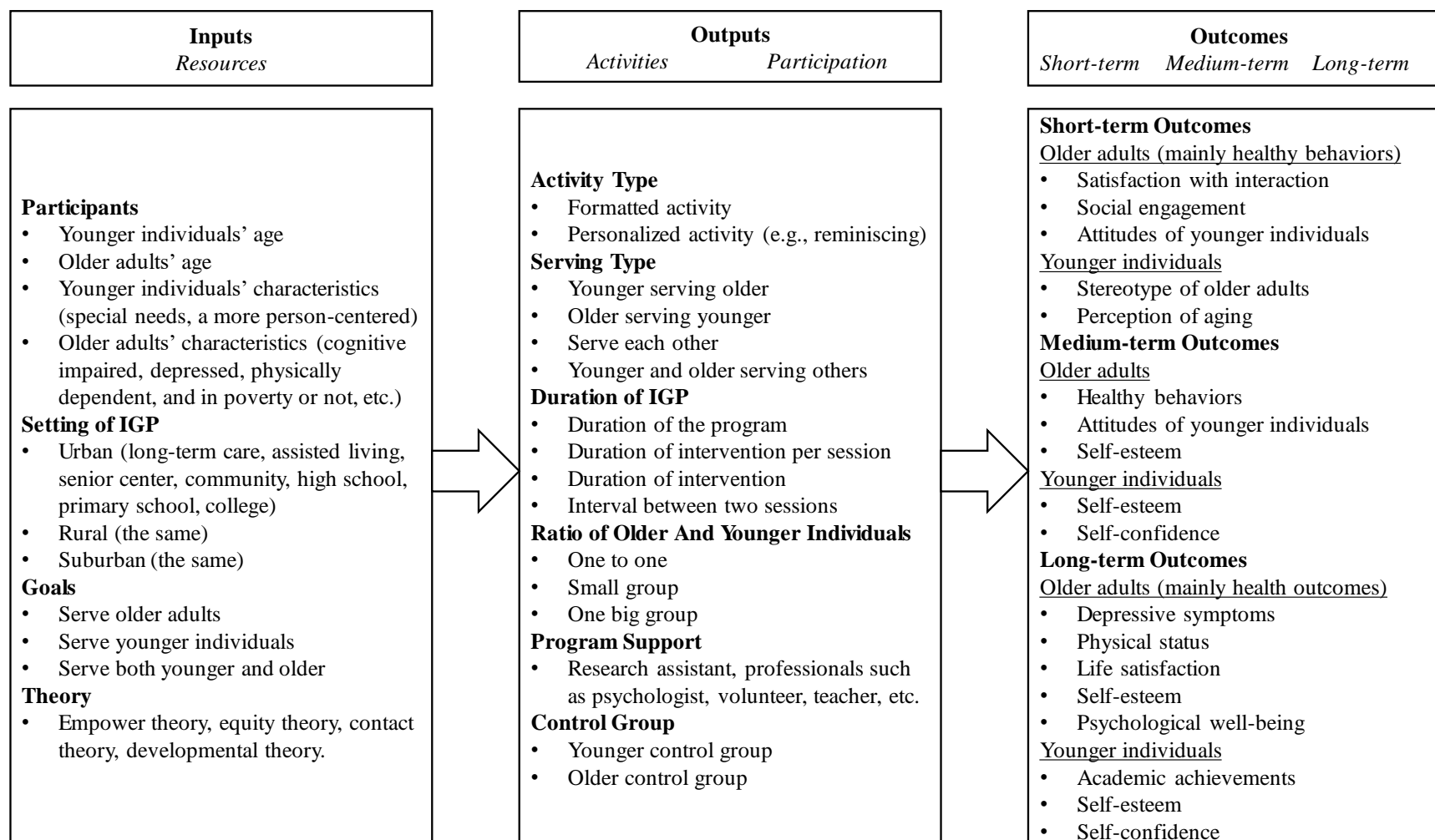


Figure 7. Logic Model Describing Suggested Elements of Intergenerational Programs (IGPs).

## TABLES

Table 1. *Basic Characteristics of All Studies Included in the Meta-analysis*

Study, First Author, & Year	Country	Source	Mean Age (OA)	Mean Age (Y)	Sample Size (OA&Y)	Outcomes
1. Gaggioli, 2014	Italy	Educational Gerontology	67.53	11	32:114	OA: Rosenberg's Self-Esteem Scale <sup>a</sup> Y: Semantic Differential Scale <sup>i</sup>
2. Murayama, 2015	Japan	Aging & Mental Health	69.10	7	80:NG	OA: Geriatric Depression Scale-Short Version-Japanese
3. Sakurai, 2016	Japan	Archives of Gerontology and Geriatrics	67.10	7	162:NG	OA: Rosenberg Self-Esteem Scale OA: Geriatric Depression Scale <sup>b</sup>
4. Belgrave, 2009	USA	Dissertation	84.75	10	26:21	OA: Rosenberg Self-Esteem Scale Y: Semantic Differential Scale
5. Posada, 2006	USA	Dissertation	84.70	2	14:13	OA: Geriatric Depression Scale
6. Hernandez, 2008	Spain	Educational Gerontology	75.00	19	101:179	OA: Geriatric Depression Scale
7. Cardona, 2002	USA	Thesis	84.20	17	13:3	OA: Rosenberg Self-Esteem Scale OA: Geriatric Depression Scale Y: Rosenberg Self-Esteem Scale <sup>j</sup> Y: Self-Efficacy Scale <sup>k</sup>
8. Meshel, 2004	USA	Educational Gerontology	75.00	12	17:63	OA: Satisfaction With Life Scale <sup>c</sup> Y: Life Satisfaction <sup>l</sup> Y: Stereotypes <sup>m</sup> Y: Attitude <sup>n</sup>
9. Smith, 2000	USA	Thesis	65.00+	7	6:NG	OA: Salamon-Conte Life Satisfaction in the Elderly Scale
10. LinOu, 2000	USA	Dissertation	86.00	4.5	15:8	OA: Salamon-Conte Life Satisfaction Scale <sup>d</sup> OA: Rosenberg Self-Esteem Scale <sup>h</sup>
11. Sue, 2014	Canada	Thesis	68.50	8	93:133	OA: Rosenberg's Self-Esteem Scale OA: Center for Epidemiologic Studies Depression Scale OA: Satisfaction With Life Survey <sup>e</sup> OA: New General Self-Efficacy Scale <sup>f</sup> Y: Attitudes to Aging and Elderly <sup>o</sup> Y: Global subscale of the Self-Esteem <sup>p</sup>
12. DeMichelis, 2015	Canada	Educational Gerontology	72.00	18	10:13	OA: Temporal Satisfaction with Life Scale <sup>c</sup>

Table 1. (Continued).

Study, First Author, & Year	Country	Source	Mean Age (OA)	Mean Age (Y)	Sample Size (OA&Y)	Outcomes
13. Yuen, 2003	USA	Physical & Occupational Therapy in Geriatrics	80.05	19	18:NG	OA: Life Satisfaction Index <sup>g</sup>
14. Gamliel, 2014	Israel	Educational Gerontology	71.50	11.5	24:27	OA: eight-item New General Self Efficacy (NGSE) questionnaire Y: General Self-Efficacy <sup>q</sup>
15. Kamei, 2011	Japan	Japan Journal of Nursing Science	76.00	8	14:7	OA: Geriatric Depression Scale OA: Health-related Quality of Life

Note. OA = older adults. Y = younger individuals. NG = not given.

**OA:**

<sup>a</sup> Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1979)

<sup>b</sup> Geriatric Depression Scale (Yesavage et al., 1983)

<sup>c</sup> The Temporal Satisfaction With Life Scale (TSWL) (Pavot, Diener, & Suh, 1998).

<sup>d</sup> The Life Satisfaction Scale (LSS; Salamon & Conte, 1998)

<sup>e</sup> Satisfaction With Life Survey (SWLS; Diener, Emmons, Larsen, & Griffin, 1985).

<sup>f</sup> New General Self-Efficacy Scale (Chen, Gully, & Eden, 2001).

<sup>g</sup> Life Satisfaction Index (Neugarten, Havighurst, & Tobin, 1961).

<sup>h</sup> The 10th article only provided scores of subscales not of the overall scale, thus self-esteem was not included in analysis.

**Y:**

<sup>i</sup> Semantic Differential Scale (Snider & Osgood, 1969)

<sup>j</sup> Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1979)

<sup>k</sup> Self-efficacy Scale (SES; Sherer et al., 1982)

<sup>l</sup> Satisfaction With Life Survey (SWLS; Diener, Emmons, Larsen, & Griffin, 1985).

<sup>m</sup> Stereotype (Fishbein & Ajzen, 1975)

<sup>n</sup> Five semantic differential scale items

<sup>o</sup> Children's Perceptions of Aging and Elderly scale (Rich, Myrick, & Campbell, 1983).

<sup>p</sup> Global subscale of the Self-Esteem Questionnaire (DuBois, Felner, Brand, Phillips, & Lease, 1996).

<sup>q</sup> New General Self Efficacy (NGSE) questionnaire (Chen et al., 2001)

Table 2. *Intergenerational Program (IGP) Characteristics*

Study	Program Support	Control Group	Interval (Weeks)	Program Duration (Weeks)	Intervention Per Session (Minutes)	Intervention (Minutes)	Ratio of Older and Younger
1	Psychologist	No	1	3	120	360	2:(6-8)
2	No	Yes	1.50	96	45+		OA: 6 to 10 members per group; Y: a class
3	No	Yes	1.50	336	45+		OA: 6 to 10 members per group; Y: a class
4	Research assistants	Yes	1	10	30	300	14:12
5	No	Yes	0.33	9	15	270	8:13
6	Trainer	Yes	1	32	50	1600	36:79 in total, then unclear in small group
7	No	Yes	0.33	3	53	473	2:1 or 3:1
8	Research assistant	Yes	1	6	60	360	1:1 or 1:2
9	Researcher	No	1	4	35	140	NG
10	No	Yes	0.50	6	30	360	1:1
11	Staff member	Yes	1	36	120	4320	1:1
12	High school teacher	No	1	3	90	270	10:13
13	No	Yes	1	7	60	180	1:1
14	No	No	1	16	120	1920	1:1
15	Older volunteers	No	1	22	180	3960	14:7

*Note.* *Interval* refers to the time between two consecutive sessions. *Program Duration* means the time the program lasts for. *Intervention per Session* signifies the contact time for every IGP meeting. *Intervention* serves as the total IGP contact time (the product of meeting times and intervention per session). *Ratio of Older and Younger* represents the ratio of older adults and younger individuals shown in each IGP contact group.

Table 3. *Intergenerational Program Characteristics and Notes from Authors*

Study	Activity Type	Serving Type	Researcher Coding Notes
1	Personalized activity	YS: Serve each other	Voluntary participants, setting: social senior center in Milan
2	Social activity	YS: Older serving younger	Data was collected at three times. I chose baseline data and the data collected at the third time.
3	Social activity	YS: Older serving younger	Participants were allowed to self-select into a treatment condition rather than being randomly assigned. Older adults 55 and older were included.
4	Social activity	YS: Serve each other	The program length was not a consecutive 10 weeks as reported by the author because of school holidays and breaks
5	Social activity	Younger serving Older	Both cognitively intact and impaired older adults were recruited and included in analysis, I only examined older adults with intact cognitive functioning
6	Social activity	YS: Younger serving older	The control group is not a real control group because it is unclear whether participants in control group interacted with younger individuals or not
7	Social activity	YS: Serve each other	The RSE alpha for younger and older are from 0.85-0.88, however, it was unclear about which value was for older adults, so I utilized the average value as the reliability of RSE for older adults.
8	Social activity	YS: Serve each other	Cronbach's alpha values for scales were different for pretest and posttest, I used the pretest alpha even the posttest alpha is bigger.
9	Social activity	YS: Serve each other	On page 82, the author said older adults met for 8 sessions while younger and older adults met only 4 sessions. It appears older adults met every week without younger individuals and met every two weeks with younger individuals
10	Social activity	YS: Serve each other	NG
11	Social activity	YS: Older serving younger	Older adults' ages are below than 60. I considered one academic year as 9 months. One older adult met with 2 younger individuals a week separately. The author referred the comparison group as a control group
12	Social activity	YS: Serve each other	The student population was diverse with representatives from South Korea, Ukraine, Moldova, Bulgaria, England, and Albania in addition to Canadian Students
13	Social activity	YS: Older serving younger	Participants assigned to the mentoring group were based on the accessibility of the LTC facilities to the students. Participants in control group were told there might not be enough students for each of them to reduce potential resentment in the control group
14	Social activity	YS: Serve each other	NG
15	Social activity	YS: Serve each other	NG

*Note.* YS signifies my name. Coding with YS in the front represents that the serving type was coded by the author of this current study instead of the original author of each study. NG signifies that the information was not given. Author notes indicates the special points from each study that is taken from each study.



Table 4. *Meta-analysis Results for Older Adults*

Construct	K	Sample size	Mean delta	Variance	$SD^1$	Lower	Upper	Weighted d	$SD^2$
Depressive Symptoms	7	477	0.26	0.27	0.52	-0.41	0.93	0.23	0.45
Life Satisfaction	6	159	0.11	0.09	0.30	-0.27	0.49	0.09	0.26
Self-esteem	6	350	0.20	0.06	0.25	-0.12	0.52	0.18	0.23
Psychological Well-being	15	625	0.37	0.25	0.50	-0.27	1.00	0.32	0.43

*Note.* K= Number of Included Studies. Mean delta = mean true effect size. Variance = True variance of effect size.  $SD^1$ = standard deviation of delta. Lower and Upper represent lower and upper bound of 80% Credibility Interval. Weighted d = sample size weighted mean effect size.  $SD^2$ =standard deviation of effect size after removing sampling error variance.

Table 5. *Correlations Among Dependent and Independent Variables*

	1	2	3	4	5	6	7	8	9	10	11	12
N of sample sizes	15	14	11	13	15	15	15	15	15	15	15	15
1.Mean Age of Y	1											
2.Mean Age of OA	-.06	1										
3.Ratio of OA&Y	.23	.35	1									
4.Intervention/Session	.08	-.73**	.15	1								
5. Program Length	-.24	-.46	.12	.39	1							
6.Interval	.04	-.77**	-.29	.48	.59*	1						
7.Number of Sessions	-.29	-.47	-.01	.25	.98**	.54*	1					
8.Control Group	-.05	.39	.14	-.59*	.25	-.12	.30	1				
9.Year	-.03	-.84**	-.33	.69**	.46	.63*	.48	-.26	1			
10.Program Support	.18	-.26	-.15	.32	-.32	.18	-.30	-.38	.12	1		
11.Sample Size-OA	-.08	-.58*	-.18	.14	.84**	.60*	.88**	.39	.53*	-.10	1	
12.Cohen's d	.64*	-.22	-.08	.26	-.17	.36	-.23	-.41	.35	.50	-.07	1

*Note.* OA = older adults. Y = younger individuals. Cohen's d signifies the effect of IGP on older adults' psychological well-being.

\* $p < 0.05$ . \*\* $p < 0.01$ .

Table 6. Comparison of (Intergenerational Program) IGP effects among Serving Types, Program Support, and Control Group

Variable Names	N	Mean	SD	ANOVA and <i>t</i> -test
Serving Types				
Serve Each Other	9	0.47	0.71	F (2) = 0.31
Younger Serving Older	2	0.37	1.18	
Older Serving Younger	4	0.15	0.30	
With or Without Control Group				
No Control Group	5	0.74	0.60	t (13) = 1.64
With Control Group	10	0.19	0.62	
With or Without Program Support				
No Program Support	7	0.03	0.58	t (13) = -2.11
With Program Support	8	0.67	0.59	

Table 7. Sample Size Weighted Regressions of Participants' Ages and Program Characteristics Predicting IGP Effect Size

<i>N=14</i>	<i>B</i>	<i>SE</i>	<i>β</i>
<b><i>Participants' Ages Predicting the effect size of IGP</i></b>			
Younger age	0.09	0.02	0.81***
Older adults' age	0.00	0.02	0.03
<i>Adjusted R<sup>2</sup></i>		0.60	
<i>F</i>		10.87 ( <i>p</i> < .01)	
<i>N=15</i>	<i>B</i>	<i>SE</i>	<i>β</i>
<b><i>Program Length and Program Support Predicting the effect size of IGP</i></b>			
Program Support	0.52	0.34	0.48
Program Length	0.00	0.00	-0.09
<i>Adjusted R<sup>2</sup></i>		0.18	
<i>F</i>		2.41 ( <i>p</i> = .13)	
<i>Note.</i> *** <i>p</i> ≤ 0.001.			

Table 8. *Exemplary Program Characteristics Benefiting Older and Younger Participants*

Study	OA: Cohen's <i>d</i>	Y: Cohen's <i>d</i>	OA: Age	Y: Age	Program Duration (week)	Time /Session (min)	Program Support	Activity Type	Ratio of OA & Y	Serving Type	Control Group
1	0.19	0.51	68	11	3	120	Yes	Reminiscence	2:(6-8)	Serve each other	No
2	-0.01		69	7	96	45+	No	Picture book reading		Older serve younger	Yes
3	0.01		67	7	336	45	No	Picture book reading		Older serve younger	Yes
4	1.12	0.45	85	10	10	30	Yes	Music	14:12	Serve each other	Yes
5	-0.47		85	2	9	15	No	Music		Younger serve older	Yes
6	1.2		75	19	32	50	Yes	Talks, excursions		Younger serve older	Yes
7	-0.20	0.82 0.94	84.2	17	3	53	No	Exercise, painting, singing	2.3:1	Serve each other	Yes
8	0.31	0.77 0.09 0.14	75	12	6	60	Yes	Talent show	1:1	Serve each other	Yes
9	0.28		65+	7	4	35	Yes	Leisure		Serve each other	No
10	-0.68		86	5	6	30	No	Leisure	1:1	Serve each other	Yes
11	0.00	-0.02 0.00	69	8	36	120	Yes	Mentor	1:1	Older serve younger	Yes
12	1.66		72	18	3	90	Yes	English class	10:13	Serve each other	No
13	0.59		80.05	19	7	60	No	Learning English	1:1	Older serve younger	Yes
14	0.98	1.20	71.5	11.5	16	120	No	Computer room activity	24:27	Serve each other	No
15	0.60		75.6	8	22	180	Yes	Leisure	14:7	Serve each other	No

*Note.* OA = older adults. Y = younger individuals. *Program Duration* refers to the number of weeks that comprised contact. *Time / Session* signifies the contact time for every IGP meeting. *Ratio of Older and Younger* represents the ratio of older adults and younger individuals in each group.

Table 9. *Difficulties in Conducting Meta-analysis for Intergenerational Programs*

Study Design Problems	Data Reporting Problems	Presenting Problems	Sample Problem
Cross-sectional design without control group	Report only means without SD	“Intergenerational program” as keywords but not appear in the paper again	Failure to provide mean age but only provide an age range
Change measure scales for pre-post data collection	Fail to provide non-significant data		
	Report only mean and SD for subscales instead of the overall scale	Mentioned an improvement of quality of life but actually did not measure quality of life	Too small of a sample size
Consider another treatment group as control group	Failure to provide the specific questions of those unfamiliar measures		
Treat experiment group with various interventions within the group	Report only post test data for the experiment group		
Random assignment is limited	Case study but reported enough information to get effect size		
Offer the control group with different intervention	Failure to report or report reliability of the measure from other authors		

Table 10. *Summary of Research Questions and Interpretations*

	<b>Analysis/information</b>	<b>Interpretation</b>
<i>Aim 1, RQ1: overall effect</i>	Meta-analysis (See Table 4)	<ul style="list-style-type: none"> <li>• IGP was effective for OA but the effect was variable</li> </ul>
<i>Aim 2, RQ 2: moderators</i>		
Age of participants	Correlation, scatterplots, & sample-weighted regression	<ul style="list-style-type: none"> <li>• OA benefited from IGP with “older” younger counterparts; 60-70 age group had the highest IGP effect; the combination of oldest OA and the youngest Y had the least IGP effect; the combination of 60-70 year-old-OA with college students had the highest IGP effect.</li> </ul>
Program support	Independent samples <i>t</i> -test	<ul style="list-style-type: none"> <li>• OA benefited from IGP with program support</li> </ul>
Program duration	Correlation and scatterplots	<ul style="list-style-type: none"> <li>• Longer programming did not necessarily benefit OA</li> </ul>
Interaction time per session	Correlation and scatterplots	<ul style="list-style-type: none"> <li>• 45 minutes to 120 minutes seemed more beneficial to OA</li> </ul>
Interval between sessions	Correlation and scatterplots	<ul style="list-style-type: none"> <li>• OA benefited more from IGP with one week’s interval and benefited less from IGP with a shorter interval</li> </ul>
Ratio of older and younger	Correlation and scatterplots	<ul style="list-style-type: none"> <li>• Ratio of older and younger was not found as related to OA’ benefits</li> </ul>
Serving type	<i>ANOVA</i>	<ul style="list-style-type: none"> <li>• OA benefited from IGP with serving each other and younger serving older types</li> </ul>
Control group	Independent samples <i>t</i> -test	<ul style="list-style-type: none"> <li>• IGP with control group had lower effect compared with IGP without control groups</li> </ul>

*Note.* OA = older adults; Y = younger individuals.

## APPENDIX A. COMPARISON OF LITERATURE SEARCH DATABASES

Table A1. Comparison of Literature Search Databases

Content	Database	Quoted Description <sup>a</sup>
Technology, medicine, social sciences, and arts and humanities	Scopus	A large abstract and citation database of peer-reviewed literature, delivers a world's research output in the fields of science
Psychology and psychology-related topics	PsycINFO	PsycINFO provides access to international literature. Nearly all records contain no evaluative summaries, and all records from 1967 to the present are indexed
Education	ERIC	The Education Resources Information Center (ERIC) is a digital library of education-related resources sponsored by the Institute of Education Sciences (IES) of the U.S. Department of Education
Multiple disciplines	Google Scholar	From one place, you can search across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations. Google Scholar helps you identify the most relevant research across the world of scholarly research.
Multiple disciplines	Web of Science	Web of Knowledge allows you to search across multiple databases. Searched databases include: Web of Science, Conference Proceedings Citation Index, BIOSIS Previews, CAB Abstracts, and Zoological Record.
Biomedicine and health	PubMed	Includes the fields of biomedicine and health, covering portions of the life sciences, behavioral sciences, chemical sciences, and bioengineering.
Multiple disciplines	EBSCO	This database covers almost all subject areas and is especially good at interdisciplinary areas. Good place to start for basic term papers. It contains full text articles for more than 2,100 journals.
Psychology and behavior	PsycARTICLE direct	Full-text of journals from the American Psychological Association, totaling about 80 journals on psychology and the behavioral sciences



Table A1. (*continued*)

Sociology and related disciplines	Sociological Abstracts	Sociological Abstracts and indexes the international literature in sociology and related disciplines in the social and behavioral sciences. The database provides abstracts of journal articles and citations to book reviews drawn from over 1,700 serials publications, and also provides abstracts of books, book chapters, dissertations, and conference papers. Records added after 1974 contain in-depth and nonevaluative abstracts of journal articles.
Multiple disciplines	Dissertations & Theses	Provides access to abstracts and citations to United States and selected foreign dissertations published since 1861. Full-text is available for dissertations from 1997 to the present. It also contains “some” full text coverage for older graduate works. Those that are not available in full-text online can be requested through Interlibrary Loan

Note. <sup>a</sup> = Quoted text from Iowa State University: <http://www.lib.iastate.edu/finddb2-startform/6023>

## **APPENDIX B. ARTICLE CODING MANUAL**

This manual is divided into four sections: basic characteristics of each article, participant information, program information, and effect size related information.

- Section one includes basic characteristics of each article: whether include in meta-analysis, whether need to ask for data, the number of unique studies, author, year of publication, source of publication, country of the study, and notes taken from each individual article.
- Section two includes participant information: mean age and SD of older and younger participants, the percent of male older and younger participants, and the race of older adults and younger participants.
- Section three depicts program characteristics: ratio of older to younger participants, serving type, program support, resource of program support, pre-posttest, control group, time interval (weeks), length of program (weeks), number of sessions, average length of intervention (minutes), total length of intervention (minutes), type of activity, setting of IGP, and intervention description.
- Section four describes information related to effect size: dependent variable, measure of dependent variable, meaning of the measurement, dependent variable-reliability, dependent variable-type, dependent variable-trait, reliability, total sample size-older adults, treatment sample size-older adults, total sample size-younger individuals, type of reliability, original effect size, type of effect size, correct effect size, artificial dichotomization of dependent variable.

### Section One: Basic Characteristics of Each Article

Variable Names	Codes/Explanations/ Special Notes
Include in meta-analysis	1=Yes, 0=No
Need to ask for data	1=Yes, 0=No
Number of unique studies	The number of studies after checking sample duplication Note: If two studies had the same program name but the sample size and mean age were quite different, then the two studies were considered as two unique studies.
Author	The last name of the first author
Year of publication	Note: For unpublished thesis and dissertation, this means the year the dissertation or thesis was presented.
Source of publication	The name of the journal on which the paper was published Note: If the paper was not published, the source of publication was then coded as thesis or dissertation.
Country of the study	The country is where IGP happened rather than the country where the author is located
Note	This variable reported the limitations and special notes in each study reported by each author

## Section Two: Participant Information

Variable Names	Code/Explanation/Special Notes
Mean age of older participants	Note: 1. When author only provided mean age of experimental group and control group individually, I averaged the ages from both groups to get mean age for the total older adult's sample. 2. When author only provided age range instead of mean age, then the mean age was the average of the age ranges. When the age range was above a certain age, it was considered as missing. 3. If the mean age is over 60 but older adults aged from over 55 years old to 90 years old, I discussed with committees and included that article.
Standard deviation of older participants' age	The standard deviation of the mean age of older participants
Mean age of younger participants	1. If authors only provided "children in primary school" or "kindergarten" as the age period of younger individuals, I used the conventional age as representative of the age period. 2. For an age range of kindergarten to primary school, I took the medium point of the age range as the mean of younger participants.
Standard deviation of younger participants' age	The standard deviation of the mean age of younger individuals
Male percent of older participants	Male percent in the total older adult's sample
Male percent of younger participants	Male percent in the total younger individual's sample
Race of older adults:	White, Black, Asian, American Indian Native, or Native Hawaiian
Race of younger individuals:	White, Black, Asian, American Indian Native, or Native Hawaiian

### Section Three: Program Information

Variable Names	Code/Explanation/Special Notes
Ratio of older to younger	The ratio of older adults and younger individuals in per group instead of the ratio of total number of older adults and total number of younger individuals in the study
Serving type	Note: 1.This coding is based on the activity type. 2.If the program is inviting older adults to read picture books to younger individuals, the serving type was then coded as older serving younger. 3.Older as mentor was coded as older serving younger even though younger individuals visit older to receive mentor.
Program support	Whether there is a facilitator to assist the interaction of younger and older during the IGP intervention. Note: If there are other participants besides older and younger individuals but they were not facilitating the interaction between younger and older participants, then those persons were not coded as program support. For example, teachers were present to help infants but were not encouraging the interaction between younger and older and researchers were observing the interaction between younger and older adults.
Resource of program support	Who provided program support? Note: A program support is not considered as another generation. For example, a psychologist who was present during children and older adults' interaction was not considered as another generation. Therefore, IGP is a two-generation activity instead of three-generation activity.
Pre-post	1=Yes. 0=No. Whether the program is a pre-post design study
Control group	Whether the study has a control group The author includes articles both with and without control groups if the study has enough information that can be used to get effect size.
Time interval (weeks)	The interval time between two consecutive sessions Note: Once every one to two weeks is considered as once per 1.5 week
Length of program (weeks)	The time the program lasts. Note: An academic year is considered as 9 months. Every month is considered as 4 weeks.
Number of meeting	The total number of meetings within program length.
Average length of intervention (minutes)	The time of every meeting
Total length of intervention (minutes)	The total intervention time and it is calculated by timing the numbers of meeting with average length of meeting together.
Type of activity	The classification of the activity content. It is also a short name of the activity

Setting of IGP	Where IGP happened. It could be rural, urban, or suburban. Among these areas, it also includes nursing home, assisted living, community, or schools.
Activity description	The detailed description of the activity

#### Section Four: Information Related to Effect Size

Variable Names	Code/Explanation/ Special Notes
Dependent variable	The name of dependent variable
Measure of dependent variable	Specific name of the measure scale used in the paper
Meaning of the measurement	What does higher score mean
Dependent variable -reliability	
Dependent variable type	1= older adults, 2=younger individual
Dependent variable-trait	Self-report or not
Reliability	The reliability from the current study and the targeted sample Note: If the author only provides a reliability range, the lower value was used If the reliability the author reported is from previous study, it is considered as missing
Total sample size-older adults	
Treatment sample size-older adults	
Total sample size-younger individuals	
Type of reliability	Alpha, test-retest, interrater
Original effect size	Effect sizes before reversing depressive symptoms
Type of effect size	Cohen's <i>d</i> , Glass's $\Delta$ , Hedges' <i>g</i> , or Morris' <i>d</i>
Correct effect size	Effect sizes after reversing depressive symptoms
Artificial dichotomization of dependent variable	1=Yes. 0=No

**APPENDIX C. IRB EXEMPTION**

**From:** Yan Su [mailto:[yansu@iastate.edu](mailto:yansu@iastate.edu)]  
**Sent:** Monday, March 6, 2017 9:12 AM  
**To:** IRB Committee [ORR] <[irb@iastate.edu](mailto:irb@iastate.edu)>  
**Subject:** IRB inquiry

Dear Sir or Madam,

My name is Yan Su, a second-year master student majoring in Gerontology and my home major is Human Development and Family Studies at Iowa State University. I am doing a meta-analysis for my thesis. I will collect published articles and theses and dissertations. I am writing to ask if I need to submit an IRB. If I need to submit an IRB, do I need to submit exempt or other types of IRB?

I am looking forward to your reply. Thank you very much!

Sincerely,  
Yan Su



On Mon, Mar 6, 2017 at 2:45 PM, IRB Committee [ORR] <[irb@iastate.edu](mailto:irb@iastate.edu)> wrote:

Yan Su,

To determine if a project needs IRB review, 2 criteria need to be met.

1. Is it research? Yes

- a. ISU considers dissertations and thesis work as research. And by your description it fits the federal definition for research.

2. Are there Human subjects?

A human subject is defined as a living individual about whom an investigator conducting research obtains (1) data through intervention or interaction with the individual, or (2) identifiable private information about whom includes a subject's opinion on a given topic.

Without knowing the nature of the publications and sources for your analysis it is difficult to know if you need oversight. For example, If these sources have identifiable information of participants or if the publications included the dataset then it would need IRB oversight.

If you do have human subjects let me know and we can discuss whether it can be classified as exempt.

Let me know if you have any more.



IOWA STATE UNIVERSITY  
BIOLOGY

Institutional Review Board (IRB) Administrator  
2420 Lincoln Way, Ames, IA 50014  
[515.294.1329](tel:5152941329)

IRB Committee [ORR]

[irb@iastate.edu](mailto:irb@iastate.edu)



[Show details](#)

[REDACTED]@iastate.edu>

2:02 PM (2 minutes ago) ☆



People (5)

IRB Committee [ORR]

irb@iastate.edu



Show details

\*\*\*

----- Forwarded message -----  
 From: IRB Committee [ORR] <irb@iastate.edu>  
 Date: Tue, Mar 14, 2017 at 9:10 AM  
 Subject: RE: IRB inquiry  
 To: "Su, Yan [HD FS]" <yansu@iastate.edu>

Yan Su,

Thank you for the well wishes. From your description, IRB approval is not needed. From my understanding, you are not interacting with human subjects or obtaining identifying information about them.

[REDACTED]  
**From:** Yan Su [mailto:yansu@iastate.edu]  
**Sent:** Monday, March 13, 2017 4:38 PM  
**To:** IRB Committee [ORR] <irb@iastate.edu>  
**Subject:** Re: IRB inquiry

[REDACTED]  
 Thank you for your detailed response.

I am collecting published articles that do not have identifiable information of participants. The professor who teaches meta-analysis told me that I would not need an IRB because I am not dealing with human subjects. I am simply coding existing information.

Therefore, I am wondering if you can clarify that whether I need IRB or not. Thank you very much!  
 I am looking forward to your reply!  
 Have a wonderful spring break!

Sincerely,  
 Yan Su

On Mon, Mar 6, 2017 at 2:45 PM, IRB Committee [ORR] <irb@iastate.edu> wrote:

[REDACTED]