Does cohort matter in the association between education, health literacy and health in the USA?

Takashi Yamashita¹* and J. Scott Brown²,³

¹Department of Sociology, University of Nevada, Las Vegas, Box 455033, 4505 S. Maryland Parkway, Las Vegas, NV 89154–5033, USA, ²Scripps Gerontology Center, Miami University, Oxford, OH 45056, USA and ³Department of Sociology and Gerontology, Miami University, Oxford, OH 45056, USA

*Corresponding author. E-mail: takashi.yamashita@unlv.edu (T.Y.); sbrow@muohio.edu (J.S.B.)

Summary

Growing empirical evidence supports the generally positive relationship between education, health literacy and health outcomes. However, little is known about cohort in this relationship. This study examined the role of cohort defined by 10-year age period in the association between educational attainment, health literacy and self-rated health. The data were obtained from the 2003 National Assessment of Adult Literacy survey restricted file. Focusing on nationally representative community-dwelling adults age 25 years and older, self-rated health was modeled as a function of health literacy, educational attainment, cohorts (defined by 10-year age periods), other demographic characteristics and socio-economic status. While the youngest cohort was positively associated with self-rated health, middle-age cohorts were more likely to have lower self-rated health (compared with the age 65 years and older cohort). Interestingly, age was no longer statistically significant after adjusting for cohort and other covariates. Recognition of possible cohort effects in education, health literacy and health should be reflected in future health literacy research and intervention programs for addressing health disparities in the USA.

Key words: social class; socio-economic status; age group; life course

INTRODUCTION

Educational attainment, or the lack thereof, is one of the ‘fundamental causes of ongoing health disparities in the USA (Link and Phelan, 1995). Unfortunately, possible interventions require resource-intensive and time-consuming structural-level (e.g. education system) changes. As an alternative approach, health literacy or a set of abilities to make better use of health-related information has received increasing attention because of its role as a mediator of the association between education and health (Nutbeam, 2000; Bennett et al., 2009; von Wagner et al., 2009; Berkman et al., 2010). Development of a health literacy-based health disparities intervention program should be an urgent task (Paasche-Orlow and Wolf, 2010; Parker and Ratzan, 2010). At the same time, better understanding the role educational attainment plays in health disparities in the USA is a necessary first step, as the meanings attached to educational attainment appear to vary among different cohorts (Thelin, 2004). As a result, this study seeks to examine possible cohort effects in the association between health literacy, educational attainment and health outcomes.

The relationship between education and health has been well-established in recent decades (Adler et al., 1994; Kawachi et al., 1997; Marmot, 2005). In general,
higher educational attainment is associated with better health outcomes over the life course (Kitagawa and Hauser, 1973; Wilkinson and Marmot, 2003; Institute of Medicine, 2009). Indeed, educational attainment is noted to be a ‘fundamental cause’ of health problems, as well-known health determinants such as economic well-being and pro-health behaviors are themselves often a function of educational attainment (Link and Phelan, 2005; World Health Organization Commission on the Social Determinants of Health, 2007). A number of competing hypotheses explain how lower educational attainment negatively influences health outcomes through fewer economic resources, poorer healthcare accessibility, greater stress, lack of social support, living in an unsafe environment and undesirable health behaviors (e.g. smoking) (Winkleby et al., 1992; Ross and Wu, 1995; Ross and Mirowsky, 1999). Such explanations are useful for informing interventions that address one of the national health goals: reduction of health disparities for individuals from diverse demographic (e.g. gender, race) and socio-economic (e.g. education, income) backgrounds in the USA.

However, while several studies support the importance of cohort in the context of education and health (Ross and Wu, 1996; Krieger et al., 1997; Yen and Moss, 1999; Lynch, 2003), the potential role of cohort in view of health literacy is not yet known. Any consideration of cohort differences illuminates several issues with education and health outcomes research. First, formal education likely has different meanings, both qualitatively and quantitatively, across different cohorts. For example, a contemporary college education experienced today by 18–25 year olds, when compared with those whose college education was completed 50 or more years ago, should be different as a result of dramatic changes over time in access to and in the structure of educational systems. Second, accessibility to college education may alter meanings of education depending on the development of and economic conditions in a society at the time of formal education. Finally, although educational attainment has been recognized in social science research as a credentialing system that enables access to social and economic privilege (Collins, 2006), a large number of studies still simply use measures of education that are either the total years of formal education completed or highest degree earned (Braveman et al., 2005). As a result, the period of time since completion of formal schooling, as well as any additional formal or informal learning experiences (e.g. life-long or life-wide learning) beyond the completion of formal schooling are typically ignored (Desjardins, 2003; Pamphilion, 2005; Wister et al., 2010). This makes some commonly used measures of formal education more appropriate for younger cohorts still in school or most recently completing schooling than for older cohorts who are temporally far removed from completion of their formal schooling. For example, years of schooling becomes a problematic measure when comparing younger to older cohorts (e.g. 11 years of schooling for a young adult today equates to being a high school drop-out, but for individuals among the oldest old may equate to being a high school graduate since the twelfth grade was not the standardized endpoint for USA secondary education until the middle of the twentieth century).

Health literacy, or a set of abilities to locate, understand and use health-related information, has received increasing attention from healthcare professionals as well as health educators (Sentell, 2003; Nielsen-Bohlman et al., 2004; Schillinger et al., 2006; Bennett et al., 2009). Health literacy reflects more ‘updated’ practical skills and capabilities that are associated with health outcomes. Such skill sets are indispensable in a society where healthcare systems and medical research are constantly evolving (St Leger, 2001; Schillinger and Davis, 2005; Askelson et al., 2011; Luk and Aslani, 2011). Additionally, one’s health needs and necessary health information are not fixed over the life course but may be uniquely different in each stage/age of life. In fact, Bennett et al. (Bennett et al., 2009) found that health literacy is a significant mediator between educational attainment and health outcomes after adjusting for covariates in a nationally representative sample of American older adults. Therefore, taking health literacy into account is sensible in health-education research because health literacy may be a more current indicator of education status as well as an indicator of learning outcomes (e.g. timely health knowledge) (Rootman and Ronson, 2005). Health literacy may fill a time gap in the relationship between formal education and health outcomes particularly for middle-age to older cohorts.

Thus, investigating cohort differences in the association between education (i.e. traditional formal education measures), health literacy and health outcomes is critical. In general, designing intervention programs is challenging for diverse populations. Ordinary public health interventions are either individual-focused (i.e. providing a tailored intervention of each person) or population-focused approaches (i.e. providing a one-size-fits-all intervention) (Rose et al., 2008). However, these approaches have known limitations including tremendous resource needs (e.g. specialists, cost, time) and cultural insensitivity (e.g. the same ‘one-size-fits-all’ intervention for young and older persons) (Smedley and Syme, 2001). While community-level (e.g. city, county)
approaches have the potential to overcome some of the shortcomings in typical public health interventions, knowledge about cohort differences would further support decision-making for design of interventions based on possible mechanisms of such differences and for efficiently allocating resources to cohorts with greater needs.

Changing formal educational attainment at a population level is particularly difficult as such modification necessitates structural-level changes in educational systems, as well as a long passage of time (World Health Organization Commission on the Social Determinants of Health, 2007; Freedman et al., 2009). In this respect, focusing on health literacy, which is a mediator of the association between education and health, is more sensible (Sentell, 2003; Zarcadoolas et al., 2006). This is particularly the case for older adults—who completed their formal education long ago—and who are generally at a greater risk of adverse health consequences. Some suggest that efforts at improving formal educational attainment may even be inappropriate as a public health strategy in view of the theoretical pathways from education to health (Andersen, 1995; Freedman and Martin, 1999). Indeed, recent attempts to reduce public health disparities through changes in education have not yet made significant impacts in developed nations (Nutbeam, 2000). Insufficient attention to cohort differences arguably has played a role in this context.

Building upon the study of Bennett et al. (Bennett et al., 2009), this study explores the possible role of cohort in the associations between health literacy, education and health outcomes measured by self-rated health using a nationally representative sample of American adults. Statistical significance of cohort measures in regression analysis would imply potential differences in the impact of health literacy and/or education on health outcomes. Given the diverse experiences of cohorts in formal education, including historical changes in education systems in the USA, as well as changes in the medical system including the transition to a consumer orientation, it is hypothesized that cohorts are independently associated with self-rated health after adjusting for established predictors including educational attainment and health literacy. Also, we expect that possible effects of cohort on self-rated health are stronger in older cohorts than younger cohorts.

**METHODS**

**Data**

Data are derived from the 2003 National Assessment of Adult Literacy (NAAL) restricted file (White and Dillow, 2005). We obtained the data license from the US Department of Education’s National Center for Education Statistics (NCES). The NAAL conducted a series of in-person interviews for a nationally representative probability sample of households and included inmates in prisons. The detailed description of the NAAL sampling strategies has been reported elsewhere (Baldi, 2009). The NAAL-restricted file contains sets of demographic, socio-economic, health and literacy skills information. Out of >18 000 (total sample) NAAL respondents, 12 930 American adults age 25 and older living in community are included in this current study after excluding the respondents under age 25 (~16.8% of total sample), missing information (~4.3% of total sample) and prison populations (~7% of total sample). Although the NAAL respondents included those aged 16 and older, we focus on 25 years and older because of the importance of completed college education. Indeed, the majority of undergraduate students are under 25 years old in the USA (U.S. Census Bureau, 2005).

**Measures**

Measures are selected according to theoretical relevance and availability in the NAAL-restricted file. Health literacy is a derived measure from 28 health-related questions of three different literacy domains including prose (e.g. searching specific information in written materials), document (e.g. filling out a job application form) and quantitative (e.g. performing numeric tasks) literacy in the NAAL. Some examples of such questions regarded information about the appropriate dose of a prescribed medication on a prescription label, the health risk/benefit from written materials of nutrients, and the benefits of health insurance plans. Sample questions can be found in the NAAL website (http://nces.ed.gov/NAAL/si_questions.asp). The NAAL employs a balanced incomplete block spiraling design (i.e. each survey participant completes ~25–30% of 152 questions). As such, while the NAAL provides the descriptive summary of health literacy assessment results in a score (0–500) for the entire sample or for sub-groups, a health literacy score estimate for individual respondents is not available. Instead, health literacy is recorded in an item response theory (IRT) theta scale, ranging from −4 to 4 (see Bennett et al., 2009 for more details).

Self-rated health is dichotomized into (1) excellent, very good and good health and (0) fair and poor health. This decision was made based on the data distribution and previous studies; self-rated health is a widely used valid measure of health status (Ferraro and Farmer, 1999; Jylhä, 2009), and dichotomization is commonly
Does cohort matter in the association between education

used (Kawachi, 1999). Age is recorded in years. Female (vs. male), Black (vs. White and others), Hispanic (vs. White and others), College education (vs. less than college education), Marital status (vs. not married), Employment status (vs. not employed) and Born in the USA (vs. not born in the USA) are dichotomous measures. Household income is recorded in 12 categories (0 ≤ $5000; 1 ≤ $5001 to $7499; 2 ≤ $7500 to $9999; ... 10 ≤ $60 000 to $74 999; 11 < $75 000 to $99 999; 12 ≥ $100 000). A series of cohort measures was created from the age/birth year data and indicated the respondents’ cohort (25–34 years; 35–44 years; 45–54 years and 55–64 years) vs. those aged 65 years and older. While most of these cohorts are defined by 10-year age-period, the group of 75 years and older (~6.7% of total sample) is included in the oldest cohort due to limited sample size.

Statistical analysis
All statistical analyses are conducted using the AM software (available at: http://am.air.org/default.asp). The AM software was developed by the American Institute of Research for analyzing large-scale complex assessment data (Cohen, 2011). Weighted summary statistics for all variables for cohorts (means and standard deviations/standard errors) are reported separately. The survey sampling weights provided in the NAAL survey data are incorporated into all analyses. Marginal maximum likelihood (MML) probit regression is used to model self-rated health as a function of education, health literacy and other covariates (Liao, 1994; DeMaris, 2004; Cohen, 2011). The AM software has the capability to run the MML probit regression model with health literacy measured in the IRT theta scale. To examine the impact of cohort on the associations between self-rated health, age and education, three different models are constructed: a model with only the age measure (no cohort variables) (Model 1), a model with only the cohort measures (no age variable) (Model 2) and a model with both age and cohort measures (Model 3). All analyses in this study strictly follow the guidelines required in the NCES data license agreement within the capability of the AM software.

RESULTS
A descriptive summary of the entire sample of respondents as well as of each age cohort is shown in Table 1. The mean age for all respondents is 48.2 years old (SD = 15.5). There are slightly more females than males in all cohorts. Almost 8 in 10 respondents (77.9%) are White or other races, not including Blacks (11.8%) and Hispanics (10.3%). The proportions of Black and Hispanic are greater in younger cohorts than older cohorts. Approximately two of three (65.7%) respondents are married, and marital status is comparable across cohorts (ranging from 60 to 69% married). Household income is also fairly consistent across cohorts except for the oldest cohort which has lower income. Employment rate is also significantly lower in the oldest cohort compared with younger cohorts. Nine in ten respondents (86.7%) were born in the USA. Across all respondents, 85.0% rate their own health as excellent, very good or good. However, there is a linear decline from the youngest (93.4%) to the oldest cohort (72.3%) in self-rated health such that respondents in the older cohorts are most likely to have lower ratings of their own health. Almost one in three (36.6%) respondents has 2 years or more of college, but again, there is a linear decline from the youngest (40.7%) to the oldest cohort (24.7%) in the likelihood of completing 2 or more years of college. Finally, as cohorts aged, the health literacy score declined; the youngest cohort has an average score of 257.4 (SE = 1.8), whereas the oldest cohort averaged 213.9 (SE = 2.0).

Probit regression results
Table 2 summarizes the final results of the three probit regression models. In Model 1, all measures with the exception of Blacks (vs. Whites and other races) are statistically significantly associated with self-rated health. Health literacy is independently and positively associated with self-rated health, after controlling for college education and the other covariates. Older age and being born in the USA are negatively associated with self-rated health. In Model 2 (which includes cohorts but excludes age), all cohorts (vs. the 65 and older cohort) are statistically significantly associated with self-rated health. However, two middle-aged cohorts (45–54 and 55–64 years old) are negatively associated with self-rated health, whereas two younger cohorts (25–34 and 35–44 years old) are positively associated with self-rated health (vs. the 65+ cohort). In Model 3 (which includes both age and the cohorts), age is not statistically significant after controlling for cohorts and other covariates, whereas health literacy and education remain statistically significant. Also, one of the cohorts (34–44 years old) is no longer significant in Model 3.

Across all three models, health literacy, college education, being female, marital status, household income and employment status are positively and statistically significantly associated with self-health ratings; birth
Table 1: Weighted summary statistics of the NAAL respondents by age cohorts\textsuperscript{a,b,c}

<table>
<thead>
<tr>
<th>Variables</th>
<th>All age (25 and older)</th>
<th>Age 25–34 (18.8%)\textsuperscript{d}</th>
<th>Age 35–44 (20.9%)\textsuperscript{d}</th>
<th>Age 45–54 (18.5%)\textsuperscript{d}</th>
<th>Age 55–64 (12.3%)\textsuperscript{d}</th>
<th>Age 65+ (14.6%)\textsuperscript{d}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-rated health (excellent, very good, good)</td>
<td>85.0%</td>
<td>93.4%</td>
<td>90.1%</td>
<td>84.5%</td>
<td>79.2%</td>
<td>72.3%</td>
</tr>
<tr>
<td>Health literacy\textsuperscript{e}</td>
<td>NA</td>
<td>257.4 (SE = 1.8)</td>
<td>253.2 (SE = 2.1)</td>
<td>249.2 (SE = 1.9)</td>
<td>242.3 (SE = 2.2)</td>
<td>213.9 (SE = 2.0)</td>
</tr>
<tr>
<td>College education\textsuperscript{f}</td>
<td>36.6%</td>
<td>40.7%</td>
<td>39.6%</td>
<td>38.6%</td>
<td>36.1%</td>
<td>24.7%</td>
</tr>
<tr>
<td>Age</td>
<td>48.2 (15.5)</td>
<td>29.7 (2.9)</td>
<td>39.5 (2.8)</td>
<td>49.2 (3.0)</td>
<td>59.1 (2.9)</td>
<td>74.1 (6.4)</td>
</tr>
<tr>
<td>Female</td>
<td>52.5%</td>
<td>51.7%</td>
<td>51.6%</td>
<td>52.3%</td>
<td>53.4%</td>
<td>54.5%</td>
</tr>
<tr>
<td>Black</td>
<td>11.8%</td>
<td>14.1%</td>
<td>13.3%</td>
<td>12.9%</td>
<td>10.1%</td>
<td>6.9%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>10.3%</td>
<td>16.6%</td>
<td>12.8%</td>
<td>8.3%</td>
<td>6.0%</td>
<td>4.9%</td>
</tr>
<tr>
<td>Marital status (married)</td>
<td>65.7%</td>
<td>61.8%</td>
<td>67.9%</td>
<td>69.4%</td>
<td>68.7%</td>
<td>60.4%</td>
</tr>
<tr>
<td>Household income\textsuperscript{g}</td>
<td>7.65 (3.1)</td>
<td>7.7 (3.1)</td>
<td>8.0 (0.4)</td>
<td>8.2 (3.2)</td>
<td>7.8 (3.1)</td>
<td>6.2 (2.9)</td>
</tr>
<tr>
<td>Employment status (employed)</td>
<td>65.5%</td>
<td>78.2%</td>
<td>80.4%</td>
<td>79.8%</td>
<td>59.3%</td>
<td>15.0%</td>
</tr>
<tr>
<td>Born in the USA</td>
<td>86.8%</td>
<td>81.2%</td>
<td>84.3%</td>
<td>88.2%</td>
<td>90.5%</td>
<td>92.8%</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Mean (SD) for continuous variables and proportions for categorical variables.
\textsuperscript{b}NAAL sample size = 12 930.
\textsuperscript{c}Individual and cluster survey weights are applied.
\textsuperscript{d}Weighted proportion given the weighted sample populations.
\textsuperscript{e}Health literacy score estimate can be computed only for sub-groups: score range 0–500 (SE).
\textsuperscript{f}Two years associate degrees or higher educational attainment.
\textsuperscript{g}One to twelve household income categories (0 ≤ $5000; 1 ≤ $5001 to $7499; 2 ≤ $7500 to $9999; ... 10 ≤ $60 000 to $74 999; 11 < $75 000 to $99 999; 12 ≥ $100 000).
status is negatively and statistically significantly associated with respondent ratings of their own health.

**DISCUSSION**

Building on the established relationship between education and health, and the recent study of Bennett et al. (Bennett et al., 2009), we examine the role of cohort in the associations between education, health literacy and self-rated health. Regression analysis of the NAAL 2003 restricted data reveals that cohorts are independently associated with self-rated health after controlling for the covariates. Indeed, age is no longer statistically significant after controlling for cohorts (see Model 3 in Table 2). Although being in the younger cohorts has a positive effect on self-rated health, being in the two middle-age cohorts (45–54 and 55–64 years old) has negative effect (vs. 65+ cohort). Overall, results suggest possible cohort effects in the relationship between education, health literacy and self-rated health although the data analyzed in this study do not allow scientifically sound separation of age and cohort effects.

The findings support our hypothesis as the cohort variables are statistically significant even after adjusting for education and health literacy. Also, one of the important findings in this study is that age is not associated with self-rated health after adjusting for cohorts and other covariates. Arguably, qualitative differences between the possible effects of actual age and cohort (e.g. normal aging process, shared life experience) in the context of education (e.g. education system changes in history), health literacy and self-rated health might exist. For example, experiences in formal education and learning opportunities after schooling might be reflected more precisely within cohort rather than by age (Thelin, 2004; Wister et al., 2010). Whereas the cross-sectional data analyzed in this study do not allow detailed investigation of historical events (e.g. the Great Depression) or causal inference, it is possible that cohort represents a set of cohort-specific earlier life experiences including formal education, healthcare access, living environment and health behaviors in addition to the effects of the aging processes (Ross and Mirowsky, 1999). Generally, older cohorts are more likely to face a greater impact of educational attainment on health due to cumulative effects of education in earlier life (Ferraro and Shippee, 2009). Yet, this study finds that two middle-age cohorts (45–54 and 55–64 years old) are more negatively associated with self-rated health compared with the 65 years and older cohort after adjusting for covariates. Although some factors such as increasing inequality in health, socio-economic well-being, stressful work environments or selection effects (i.e. older adults as survivors) could explain the negative effects of middle-age cohorts, further study with longitudinal data is warranted.

With regard to the association between education, health literacy and health, the findings about health literacy being an independent predictor of self-rated health are consistent with previous national studies (Gazmararian et al., 1999; Sentell and Halpin, 2006; Bennett et al., 2009). As such, one may argue that health literacy is not simply an outcome of formal education (Wilson et al., 2010). Rather, it appears to be a more complex construct reflecting quality/type of formal education, learning after schooling, practical knowledge/skills about health, and presumably, individual abilities to learn new information (Schillinger and Davis, 2005; Nutbeam, 2008). While formal educational attainment is generally difficult to alter, health literacy is more feasible to improve (Zarcaoolas et al., 2006). Therefore, the findings in this study suggest cohort-specific health literacy education as a potential intervention strategy targeting adult populations. Furthermore, early intervention of limited health literacy for middle-aged cohorts may benefit their health outcomes in later life.

This study makes several contributions. First, the empirical analysis of the NAAL-restricted data demonstrates the possible existence of cohort effects to the literature regarding education, health literacy and health outcomes in the lack of nationally representative longitudinal data with sound health literacy measures (White et al., 2008; Bennett et al., 2009). Second, this study focuses on the understudied area of possible cohort effects although the key determinant of health disparities, formal educational attainment, is strongly associated with cohorts. Third, the results of this study are suggestive of several future research areas including separation of age-period-cohort effects in the association between education, health literacy and health in current middle-age cohorts, impact of major historical events (e.g. the G.I. bill, the Great Depression) and relationships between types/qualities of education and health literacy. Finally, the findings from this study highlight the need for large-scale health literacy data collection as future research will require longitudinal data. While the NAAL-restricted file is one of few nationally representative data sets with the sound measures of health literacy, the accessibility and usability are limited due to the strict data security guidelines [e.g. a dedicated room and non-networked computer are required; major commercial statistical packages cannot take the complex sampling design and rather unique measurements (i.e. IRT) into account].

---

*Does cohort matter in the association between education*
This study was not without limitations. The cross-sectional data used here do not allow for causal inference about the findings. Furthermore, it is not possible to examine within-cohort differences (e.g. all 65 years and older respondents are in the same cohort) due to sample size limitations and respondents’ personal privacy protection as required by the NAAL-restricted data guidelines (see the NAAL website for more details: http://nces.ed.gov/pubsearch/pubsinfo.asp?Pubid=2007493).

Additionally, while an ongoing discussion suggests the importance of education measurement (Braveman et al., 2005), detailed information about educational attainment such as quality of schools and subject of study are not available in the NAAL-restricted file. Another limitation to the current study is possible omitted variables (e.g. living environment, social support, healthcare accessibility) suggested in Rootman and Ronson’s (Rootman and Ronson’s, 2005) theoretical framework. Though not available in NAAL, these measures are important additions needed in ongoing investigation of health literacy. Therefore, the interpretation of findings in this study may be only applicable at the theoretical level. Finally, the data analyses conducted according to the NAAL-restricted data guidelines do not allow empirically separating age- and cohort-effects. As such, the possible cohort effects identified in this study needs verification with nationally representative longitudinal data with more scientifically rigorous methods (e.g. Yang and Land, 2008).

In conclusion, building upon the well-established relationship between education, health literacy and health outcomes, this study provides a series of empirical evidence suggestive of possible cohort effects. Another interesting finding is that middle-aged cohorts in the NAAL 2003 are more likely to report lower self-rated health compared with the older cohorts after adjusting for education, health literacy and other covariates. Additionally, the consistently positive effects of health literacy on self-rated health across all cohorts add further population-level empirical evidence to the literature and support health literacy-based health disparities interventions (Bennett et al., 2009). Possible cohort effects on health literacy, education and health outcomes should be verified with longitudinal data and be extensively discussed in future studies.

References


