

Bangor University

MASTER OF PHILOSOPHY

Religion and health: a systematic review

Quigley, Catherine Francis

Award date:
2003

Awarding institution:
University of Wales, Bangor

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 23. Jul. 2023

**RELIGION AND HEALTH: A SYSTEMATIC
REVIEW**

I'W DDEFNYDDIO YN Y
LLYFRGELL YN UNIG

TO BE CONSULTED IN THE
LIBRARY ONLY



RELIGION AND HEALTH: A SYSTEMATIC REVIEW

by

CATHERINE FRANCIS QUIGLEY

of

UNIVERSITY OF WALES, BANGOR

**A dissertation submitted for the degree of
Master of Philosophy in the University of Wales**

September 2003

Introduction	1
1 Religious Denomination and Mental Health	16
Introduction	16
Depression	17
<i>Weak Publications (Rating 1-4)</i>	19
<i>Average Publications (Rating 5-6)</i>	22
<i>Good Publications (Rating 7-8)</i>	26
<i>Excellent Publications (Rating 9-10)</i>	33
Anxiety	44
<i>Weak Publications (Rating 1-4)</i>	45
<i>Average Publications (Rating 5-6)</i>	48
<i>Good Publications (Rating 7-8)</i>	50
Subjective Well-being	53
<i>Weak Publications (Rating 1-4)</i>	54
<i>Average Publications (Rating 5-6)</i>	55
<i>Good Publications (Rating 7-8)</i>	58
<i>Excellent Publications (Rating 9-10)</i>	61
Conclusion	70
2 Religious Denomination and Physical Health	74
Introduction	74
Cancer	74
<i>Weak Publications (Rating 1-4)</i>	76
<i>Average Publications (Rating 5-6)</i>	80
<i>Good Publications (Rating 7-8)</i>	84
<i>Excellent Publications (Rating 9-10)</i>	89
Hypertension	93
<i>Average Publications (Rating 5-6)</i>	95
<i>Good Publications (Rating 7-8)</i>	97
<i>Excellent Publications (Rating 9-10)</i>	98
Heart Disease	101
<i>Weak Publications (Rating 1-4)</i>	105
<i>Average Publications (Rating 5-6)</i>	107
<i>Good Publications (Rating 7-8)</i>	108
<i>Excellent Publications (Rating 9-10)</i>	110
All-cause Mortality	113
<i>Weak Publications (Rating 1-4)</i>	114
<i>Average Publications (Rating 5-6)</i>	116
<i>Good Publications (Rating 7-8)</i>	119
<i>Excellent Publications (9-10)</i>	122
Conclusion	124
3 Public Religious Behaviours and Mental Health	129
Introduction	129
Depression	130
<i>Weak Publications (Rating 1-4)</i>	130
<i>Average Publications (Rating 5-6)</i>	132
<i>Good Publications (Rating 7-8)</i>	137
<i>Excellent Publications (Rating 9-10)</i>	144

Anxiety	152
<i>Weak Publications (Rating 1-4)</i>	153
<i>Average Publications (Rating 5-6)</i>	154
<i>Good Publications (Rating 7-8)</i>	159
<i>Excellent Publications (Rating 9-10)</i>	161
Subjective Well-being	164
<i>Weak Publications (Rating 1-4)</i>	166
<i>Average Publications (Rating 5-6)</i>	167
<i>Good Publications (Rating 7-8)</i>	171
<i>Excellent Publications (Rating 9-10)</i>	180
Conclusion	187
4 Public Religious Behaviours and Physical Health	190
Introduction	190
Hypertension	191
<i>Average Publications (Rating 5-6)</i>	192
<i>Good Publications (Rating 7-8)</i>	193
Heart Disease	196
<i>Good Publications (Rating 7-8)</i>	197
<i>Excellent Publications (Rating 9-10)</i>	197
All-cause Mortality	198
<i>Good Publications (Rating 7-8)</i>	200
<i>Excellent Publications (9-10)</i>	204
Conclusion	219
5 Private Religious Behaviours and Mental Health	223
Introduction	223
Depression	224
<i>Average Publications (Rating 5-6)</i>	224
<i>Good Publications (Rating 7-8)</i>	227
<i>Excellent Publications (Rating 9-10)</i>	231
Anxiety	236
<i>Weak Publications (Rating 1-4)</i>	237
<i>Average Publications (Rating 5-6)</i>	237
<i>Excellent Publications (Rating 9-10)</i>	239
Subjective Well-being	242
<i>Weak Publications (Rating 1-4)</i>	242
<i>Average Publications (Rating 5-6)</i>	243
<i>Good Publications (Rating 7-8)</i>	244
<i>Excellent Publications (Rating 9-10)</i>	249
Conclusion	251
6 Private Religious Behaviours and Physical Health	255
Introduction	255
All-cause Mortality	256
<i>Excellent Publications (Rating 9-10)</i>	256
Heart Disease	259
<i>Excellent Publications (Rating 9-10)</i>	259
Hypertension	261
<i>Average Publications (Rating 5-6)</i>	261

<i>Excellent Publications (Rating 9-10)</i>	262
Conclusion	264
7 Psychometric Measures and Mental Health	266
Introduction	266
Depression	268
<i>Weak Publications (Rating 1-4)</i>	269
<i>Average Publications (Rating 5-6)</i>	270
<i>Good Publications (Rating 7-8)</i>	275
Anxiety	276
<i>Weak Publications (Rating 1-4)</i>	277
<i>Average Publications (Rating 5-6)</i>	279
<i>Good Publications (Rating 7-8)</i>	281
Subjective Well-being	283
<i>Weak Publications (Rating 1-4)</i>	283
<i>Average Publications (Rating 5-6)</i>	284
<i>Good Publications (Rating 7-8)</i>	286
Conclusion	288
8 Conclusion	290
References	320

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the Revd Professor Leslie J. Francis, for his expert supervision, guidance, and inspiration, throughout the year, which I have been under his supervision. I would also like to express my sincere thanks to Dr Mandy Robbins for her all her support and help.

I would like to thank Dr Christopher Alan Lewis for his continuing help and support, throughout the year. I would also like to show appreciation to all the members of the Welsh National Centre for Religious Education, including Mike Fearn, Charlotte Craig, Anna Halsall, Douglas Turton, and Carol Roberts for their assistance and encouragement.

I am also very grateful to my family and friends, who have provided me with invaluable support, in particular Vincent, Mary and Vincent Quigley.

Catherine Francis Quigley

University of Wales, Bangor

September 2003

SUMMARY

The purpose of the present investigation is to examine the relationship between religion and health, both mental and physical, by conducting a systematic review. The relationship between the two variables is explored by splitting the various ways in which both religion and health is measured. It is recognized that researchers generally measure religion by four different traditions. Religion is measured by asking respondents to indicate their religious denomination, how often they participate in public religious activities, how often they participate in private religious activities, and more recently, through multi-dimensional motivational measures. Similarly, the concepts of mental health and physical health can be split into separate conditions. For instance, mental health can include both negative constructs (i.e., anxiety and depression) and the positive constructs (i.e., subjective well-being), whereas; physical health can include any physical illness (i.e., cancer, hypertension, heart disease, and mortality).

The areas outlined above are examined separately and assessed using an objective and systematic strategy, to deduce whether or not a relationship potentially exists. Over the years a large numbers of researchers have conducted reviews on the area, with the most recent review, conducted by Koenig, McCullough, and Larson (2001) being seen to be the most extensive and comprehensive. The present thesis examines the research cited by Koenig, McCullough, and Larson (2001) on the areas highlighted above and perform a systematic review utilizing A-Levels-of-Evidence strategy. A-Levels-of-Evidence approach has been utilized recently, i.e., Powell, Shahabi, and Thoresen, (2003), on facets of physical health. A findings from the systematic review process are discussed.

Introduction

The debate between religion and science has had a complex past, with discussion of the relationship between religion and health, both mental and physical, being no exception. The association between the two variables has been looked upon with pessimism by some researchers (e.g., Ellis, 1983) and with optimism by others (e.g., Rizzuto, 1979). Research performed in the late 1960s and in the early 1970s, recognised that religion had salutary effects on both mental health (e.g., Cooley and Hutton, 1965; Edwards and Klemmack, 1973) and physical health (e.g., Friedman and Hellerstein, 1968; Graham, Kaplan, Cornoni-Huntley, James, Becker, Hames, and Heyden, 1978). Consequently, over the years, large amounts of research has been undertaken in this area. This research has led to a plethora of publications examining the relationship between religion and health, with the majority presenting a positive picture (e.g., Koenig, 2000). Leading from the collection of studies concerned with this area, a number of researchers have conducted extensive reviews. The general conclusion stated from these reviews is that religion, in a broad sense, represents a protective factor that offers a small, but significant, primary-preventative effect against poor mental health (e.g., Bergin, 1983; Gartner, Larson and Allen, 1991) and physical health (e.g., Levin and Schiller, 1987; Levin and Vanderpool, 1989). In a recent review of the research literature, undertaken by Koenig, McCullough, and Larson (2001), this general assertion was validated.

The interest in the positive effect that religion has on health emerged due to the growing recognition that the types of health problems encountered by persons from Western societies are changing from acute diseases (e.g., influenza, tuberculosis,

diphtheria, and pneumonia), to more debilitating chronic illnesses (e.g., cancer, heart disease, immune system dysfunction, and cerebrovascular disease).

Traditionally, the biomedical model was the dominant explanation accepted by the scientific profession of what caused illness. Within this view of biomedicine, abnormalities that caused disease could not be explained by behavioural aberrations; health and illness were seen as qualitatively different, that is the individual was either healthy or ill; there was no continuum between the two. However, this reductionist viewpoint offered by the biomedical model has been criticised by Engel (1977) as being too simplistic. Consequently, Engel (1977) suggested that the medical model needed to be broadened to include the psychosocial factors that affect individuals and society. This newer model favoured by Engel (1977) was based on the theory that nothing existed in isolation; it considered health and illness to represent a hierarchy of systems and emphasised the importance of multiple levels of analysis from the microlevel (e.g., cell pathology) to the macrolevel (e.g., cultural values), these systems were considered important and interrelated.

As a result, it is now a commonly held view that health is a result of mutually interdependent physiological, socio-economic, and psychological components that exist in a state of equilibrium. Conversely, illness is referred to as the breakdown of this equilibrium within the individual or community.

For decades, the medical professions have observed how individuals in modern Western society are living longer. However, this longevity is typically accompanied by chronic illnesses that are disabling and very expensive to treat. As a consequence,

researchers have been led to examine the role of a number of variables associated with health. One such variable that has received a great deal of attention over recent years is that of religion. It has been reported that religion plays an important role in the lives of millions of people worldwide. It has been estimated that 80% of the world's population identifying themselves as adherents of one of the world religious traditions (Bernstein, Calhoun, Cegielski, Letham, Shephard, Sparks, and Tomchuck, 1995). This finding highlights the importance of religion as a variable for investigation.

The relevance of the relationship between religion and health, both physical and mental, is an area that has been characterised by much controversy over recent years. Empirical studies generally reveal a salutary relationship between the two (e.g., Colantonio, Kasl, and Ostfeld, 1992; Idler and Kasl, 1992; Koenig, 1995; Nelson, 1989).

The religious establishment has regulated and controlled science, medicine and health care (Koenig, McCullough, and Larson, 2001). In Western culture, according to Kuhn (1988), the Church was the official body that dealt with the issue of medical licenses to physicians in the twelfth century, with monks and priests tending to be practitioners. The nursing profession also emerged originally from the religious orders devoted to caring for the sick. These religious orders were also responsible for building and staffing the first medical hospital almost 1,700 years ago. Many of the first hospitals that provided compassionate care to the mentally ill were built and managed by religious groups such as the Quakers (Koenig, McCullough, and Larson, 2001). Over the past 500 years, the Church's power over the medical and health care arena declined and the influence of medical science increased. Consequently, the

medical profession has been able to disentangle itself from the control and dominance of the Church's jurisdiction.

Nonetheless, this schism has begun to mend and hence resurrect a renewed interest into the religion-health association. As a result, the discussion of how religion can help to improve health and health care was initiated, prompting investigators to examine the links between religion and health. Potentially the association between religion and health can be exploited for purposes of health promotion and disease prevention.

Psychologists have debated the positive and negative effects of religious experience. For instance, Pargament (2002) states that throughout the nineteenth century psychologists have taken both sides in the debate ranging from Skinner's view that it is punitive and exploitative to Erikson's belief that it offers a balance of wisdom and maturity. Recently, in terms of the literature, numerous authors have claimed that there is a surplus of articles supporting the contention of how religious involvement is associated with positive health effects. For instance, Koenig (2000) asserts that there are more than 850 articles on religious involvement and mental health, with more than two-thirds showing that religiously active people are better advantaged with regard to mental health. Koenig (2000) also claims that there are more than 350 articles on religious involvement and physical health, with over half showing an advantage to religiously active people. An earlier claim asserts the existence of 325 articles assessing the relationship between religiousness and physical health, with 75% showing the benefits of religious involvement (Levin and Schiller, 1987). Regardless of which review is being examined, the general opinion is that religiousness, however

operationalised, has a positive influence on the mental health and the physical health of an individual.

Koenig, McCullough, and Larson (2001) conducted one of the most extensive and comprehensive reviews in the discipline. These researchers commented on 1,200 articles concerned with the associations between religion and health. They reported that, within the literature, the evidence points to the protective effects that religiosity asserts on both the mental aspects of health and the physical entities in health. At the end of the review, Koenig, McCullough, and Larson (2001) conclude by stating that

‘In the vast majority of the cross-sectional studies and the prospective cohort studies we identified, religious beliefs and practices rooted within established religious traditions were found to be consistently associated with better health and predicted better health over time’ (p. 591).

In addition to reviewing the literature, these authors rated each of the studies, based on the overall study design, sampling method, quality of religious measure, quality of statistical analysis, interpretation of results, and discussion with in the context of existing literature. The rating scale, employed by Koenig, McCullough, and Larson (2001), ranged from one to ten (1= very poor, 10= excellent). However, a few researchers (e.g., Sloan and Bagiella, 2001, 2002) refute the conclusions drawn by Koenig, McCullough, and Larson (2001), and argue that the relationship between religion and health is of doubtful value. Sloan, Bagiella, and Powell (1999) have argued that the scientific studies conducted in the last half-century that have documented the positive effects of religion on health are hopelessly flawed and of no scientific value. The authors criticise the potential inclusion of spiritual and religious interventions in medical practice on ethncal grounds and dismiss claims that the ‘wall of separation’ between medicine and religion will be torn down, with assertions that

'The medicine of the future is going to be prayer and Prozac' (Matthews and Larson, 1997, p.85).

Sloan, Bagiella, and Powell (1999) request that the medical community should view the relationship between religion and health with caution, as the research is fraught with methodological limitations. However, it should be noted that the article in which the authors present their claims, review only 24 of the approximately 325 studies of religion and physical health. In addition, the researchers do not mention the studies that Koenig, McCullough, and Larson (2001) consider to be high quality and do not mention any of the 900 plus studies on mental health. The neglect of these key areas is not rectified in subsequent articles written by Sloan and Bagiella (2001, 2002).

Some reviewers, such as Miller and Thoresen, (2003) and Powell, Shahabi, and Thoresen, (2003) have suggested that there are associations between the two variables. Subsequently, they performed a systematic review concerning the religiousness concept in relation to facets of physical health. These authors have developed a systematic approach by adapting the Cochrane method (Miller and Thoresen, 2003). The systematic strategy was termed A-Levels-of-Evidence approach, and it was utilised to assess the quality of research in the area of religion and physical health. A-Levels-of Evidence approach establishes a well-defined set of criteria that warrants inclusion or exclusion of the research. Subsequently, the included research is placed into a category, which is set out under distinct guidelines, that describe the level of adequacy or inadequacy of the research to the area under investigation (Miller and Thoresen, 2003; Powell, Shahabi, and Thoresen, 2003).

A-Levels-of Evidence strategy was adopted by empirical researchers to assess the quality of research within the religion, spirituality and physical health arena (Miller and Thoresen, 2003; Powell, Shahabi, and Thoresen, 2003). A-Levels-of-Evidence approach was chosen over a meta-analysis because it provides a focused critique of studies, minimises bias, and is thus responsive to concerns about the quality of past reviews. Powell, Shahabi, and Thoresen (2003) subjected the research in the area of physical health to a set of strict criteria. Conclusions declared from the systematic review revealed that

‘Religion and/or spirituality could have an impact on physical health as a protective resource that prevents the development of disease in healthy people...Evidence is strongest and most consistent for a protective effect in healthy people, and this support centers largely on the hypothesis that church/service attendance protects against death’ (p. 48).

The reviewers continue to conclude that

‘Thus, we tend to agree with Sloan, Bagiella, and Powell (1999) that past reviews may tend to be over optimistic, but we tend to clearly disagree with their conclusion that suggestions that religious activity will promote health are unwarranted. Rather, we think that this conclusion is premature and that the intriguing evidence to date warrants continued and careful investigation’ (p. 50).

A-Levels-of Evidence strategy sets well-defined criteria that allows for the inclusion or exclusion of research. Research that has been excluded is placed in the category of C (inconclusive) and is eliminated from the review as their design makes it impossible to rule out bias, confounding, or chance as alternative explanations for results. To be placed in the C category the empirical research must include any of the following criteria: (a) make no attempt to control for any potential confounders; (b) have conducted a cross-sectional design; (c) provide inadequate measurement of the religion or health variable; (d) contain no statistical analyses; and (e) employed the

same cohort on which earlier reports have been based. Subsequently, if the study did not suffer these flaws they proceeded to the next stage for evaluation.

Powell, Shahabi, and Powell (2003) suggested that each of the included studies be rated for their methodological strength in two ways. The first rating (mediated model) should evaluate the impact of religion on health regardless of whether or not the relationship was mediated by established risk/protective factors (include four of the five strategies below). The second rating (independent model) should employ in its evaluation the same criteria for the first rating, but include an additional component that mediates for protective/risk factors (i.e., social support, healthy lifestyle behaviours). According to the authors, the following strategy should be adhered to in evaluating the studies: (a) adequacy of control for confounders; (b) precise measurement of religion or of covariates; (c) not failing to control for multiple tests; (d) not to perform post hoc observations in a subgroup; and (e) adequacy of control for established protective factors. If the study meets all of the above criteria, it receives a rating A (conclusive); if the study fails to meet all of the above criteria, it receives a rating of B (generally sound).

After the studies have received their rating category, the strength of the evidence for any hypothesis is judged using a scale. Powell, Shahabi, and Thoresen (2003) suggest and subsequently utilised the following scale. Evidence was considered to be *persuasive* if there were at least three supportive A studies or a mixture of five supportive A and B studies. Evidence was considered *reasonable* if there were two supportive A studies or a mixture of three or four supportive A and B studies. A hypothesis was judged as having *some* support if there was one supportive A study or

at least two supportive B studies. In cases where repeated tests of a hypothesis resulted in repeated failures to support it, the evidence was described as having *consistent failures*. If the research base was too small to fall into any of the above categories, the evidence was described as *inadequate*.

The studies that will be used for evaluation under A-Levels-of-Evidence approach for the purposes of the present thesis are the publications that have been cited by Koenig, McCullough, and Larson (2001).

The present thesis will utilise the research cited in the most recent review conducted by Koenig, McCullough, and Larson (2001) to assess whether claims surrounding the association are justified or not. The present thesis will examine the literatures on facets of mental health, such as anxiety, depression, and subjective well-being, and also components of physical health such as, all-cause mortality, cancer risk and mortality, heart disease, and hypertension. The reasoning for selecting these aspects of mental and physical health is that the areas are ones that have been frequently tested; hence, there is an abundance of published research. Consequently, the present review aims to re-examine the studies, which Koenig, McCullough, and Larson (2001) rate as the 'best' in each of their sections. The review will focus on the highest rated studies in each of the mental and physical health sections and utilise the criteria set out by the A-Levels-of Evidence approach to assess whether the claims of the protective nature of religiousness are warranted or not.

Chapter one will begin to disassemble the review conducted by Koenig, McCullough, and Larson (2001) by assessing the research on religious denomination and mental

health. The section will commence by examining the empirical literatures on the relationship between denomination and anxiety, proceeding with denomination and depression, and finally concluding with denomination and subjective well-being. The reasoning behind assessing denomination as a measure of religiousness is that the preponderance of empirical research within this domain frequently utilises religious denomination as a measure of an individuals' religiosity. In examining the mental health measures of anxiety, depression, subjective well-being allows for a plethora of research articles to be investigated and subsequently evaluated by the systematic review process (A-Levels-of-Evidence approach). Chapter one will examine each of the mental health variables separately, breaking each section into subsections according to the rating category specified by Koenig, McCullough, and Larson (2001). After the empirical studies have been investigated, the studies will be rated by utilising the evaluation strategy of the systematic review process. Subsequently, a short concluding paragraph will be offered, which sums up the evidence from the section examining each mental health variable.

The topic that will be discussed in chapter two is the relationship between religious denomination and physical health. The chapter will be divided into subsections consisting of all-cause mortality, cancer risk and mortality, heart disease, and hypertension. The empirical research examining the relationship will be assessed within the ratings assigned by Koenig, McCullough, and Larson (2001). Utilising denomination as a measure of religiousness and examining various facets of physical health allows for a substantial amount of empirical research conducted within this area that frequently utilises religious denomination as a measure of an individuals' religiosity alongside these physical health topics. Examining these two variables

alongside one another allows for a plethora of research to be investigated and subsequently evaluated by the systematic review process (A-Levels-of-Evidence approach). After the empirical studies have been investigated within these subsections, the studies will be rated by utilising the evaluation strategy of the systematic review process. Subsequently, a short concluding paragraph will be offered, which sums up the evidence from the section examining each physical health variable.

Chapter three will begin to separate the review conducted by Koenig, McCullough, and Larson (2001) by assessing the research on public measures of religiosity, in particular religious service attendance and mental health. The section will commence by examining the empirical literatures on the relationship between these public measures of religiousness and anxiety, proceeding with depression, and finally concluding with subjective well-being. Research assessing public measures of religiosity in conjunction with mental health provides a surplus of research articles to examine; hence, offering a vast number of empirical research to be evaluated by the systematic review process (A-Levels-of-Evidence approach). Chapter three will examine each of the mental health variables separately, breaking each section into subsections according to the rating category specified by Koenig, McCullough, and Larson (2001). After the empirical studies have been investigated, the studies will be rated by utilising the evaluation strategy of the systematic review process. Subsequently, a short concluding paragraph will be offered, which sums up the evidence from the section examining each mental health variable.

The purpose of chapter four is to examine public measures of religious behaviour religious services and its association with several facets of physical health (i.e., all-cause mortality, cancer risk and mortality, heart disease, and hypertension). Similarly, with the previous chapter investigating such religious measures, the preponderance of empirical research conducted into public measures of religious behaviour alongside various aspects of physical health is plentiful. Each physical health topic will be explored separately with the cited publications being investigated within their ranking category. The publications within the ranking subsections will be evaluated under A-Levels-of-Evidence strategy, for which they will receive a rating accordingly. Subsequently, there will be a concluding paragraph that sums up the evidence provided by the publications to postulate if the hypothesis stated is validated.

The theme that will be explored in chapter five is the relationship between private measures of religiosity and various facets of mental health. The chapter will begin with the examination of private religious behaviours, such as private prayer, and anxiety, proceeding with depression, and concluding with subjective well-being. The chapter will be separated into sections according to the mental health area that will be investigated. Additionally, each section will be further separated into subsections according to the rankings that the publications fall into. Research assessing private religious behaviours in conjunction with mental health provides a plethora of research articles to be examined and evaluated by the systematic review process (A-Levels-of-Evidence approach). After the empirical studies have been investigated, the studies will be rated by utilising the evaluation strategy of the systematic review process. Subsequently, a short concluding paragraph will be offered, which sums up the evidence from the section examining each mental health variable.

Comparable to the previous chapter, the next chapter (chapter six) will evaluate the association between private religious behaviours, such as prayer and physical health. The occurrence of published studies examining this domain is not as vast as the previous areas. Therefore, not all areas of physical health examined in the previous chapters will be examined due to insufficient research conducted. Therefore, the sections that will be included are all-cause mortality, heart disease, and hypertension. The empirical research will be evaluated by the systematic review process (A-Levels-of-Evidence approach); subsequently, they will receive a rating that will classify the research into one of three categories. Chapter six will examine each of the physical health variables separately, breaking each section into subsections according to the rating category specified by Koenig, McCullough, and Larson (2001). Subsequently, a short concluding paragraph will be offered, which sums up the evidence from the section examining each mental health variable.

Chapter seven focuses on the association between psychometric measures of religiosity (i.e., Allport's Intrinsic/Extrinsic Orientations; Allport and Ross, 1967) and mental health. Although, the existence of publications looking at mental health is adequate the quantity of published articles assessing physical health is inadequate. The first area that will be examined is the relationship of psychometric measures of religiosity, particularly religious motivation, and mental health. The chapter will be broken down into the following sections consisting of anxiety, depression, and subjective well-being. Additionally, the sections will be further divided into subsections depending on the rating the publications received by Koenig,

McCullough, and Larson (2001). The studies in each subsection will be evaluated in terms of the strategy assigned by A-Levels-of-Evidence approach, to deduce if the strength of the evidence allows for the judgement of the hypotheses.

The contents in chapter nine will be to aid in the understanding of the relationship and the mechanisms of the association between the variables. Chapter nine will begin with reiterating the findings of the previous chapter in relation to both Koenig, McCullough, and Larson's (2001) findings and the findings disclosed by the systematic review. The chapter will follow by investigating the limitations that are inherent in the research to date on religion and health, both mental and physical in terms of methodological and conceptual limitations. Chapter nine will offer a conclusion surrounding the relationships between the variables studied in the previous chapters of the literature review and will generate possible avenues for future research.

The purpose of this general introduction was to provide an overview of the relationship between religion and health. It can be appreciated that there are a surplus of issues that need to be addressed in understanding and explaining this association. Additionally, the aim of this present chapter was to aid in the understanding of the how the topics will be broken down and evaluated under an objective and systematic method.

Therefore, the intention of the subsequent chapters is to explore the relationship between religion and health more thoroughly by utilising the systematic review

criteria for the studies that Koenig, McCullough, and Larson (2001) deem to be the 'best' in the area and assess any limitations with previous research.

1 Religious Denomination and Mental Health

Introduction.

The principal theme to be explored within this chapter is the relationship between religious denomination and mental health. The review conducted by Koenig, McCullough, and Larson (2001) extensively assesses this theme with respect to a range of mental health issues. Hence, the intention of this chapter is to present their findings and to re-evaluate the studies that the reviewers have cited within this domain. Consequently, this chapter will critically evaluate the research conducted on mental health (anxiety, depression, and well-being). The empirical research that will be examined will be evaluated under A-Levels-of-Evidence strategy in order to observe whether the relationship between the two variables are justified as declared in Koenig, McCullough, and Larson (2001) most recent review and whether there is enough evidence presented from the empirical research to communicate if a potential relationship is likely to exist.

The existence of a relationship will be explored by utilising A-Levels-of-Evidence strategy adopted by Powell, Shahabi, and Thoresen (2003). Utilising this approach allows for an objective and systematic evaluation of the empirical research cited by Koenig, McCullough, and Larson (2001) and for an assessment of the strength of the evidence for the hypothesis (i.e., is there enough evidence to postulate if religious denomination is related to mental health).

Depression

Reviews assessing the relationship between religious denomination and depression report that the preponderance of the evidence suggests that certain denominations are at an elevated risk of suffering from depression (e.g., McCullough and Larson, 1999). McCullough and Larson (1999) reviewed the literature on approximately 80 studies, which examined the association of religious denomination or involvement alongside depressive symptoms or depressive disorders. The reviewers concluded that

‘Jews and people not affiliated with a religion were at an elevated risk of depression and depressive symptoms. These particular findings appear to be quite robust and well replicated across many data sets’ (McCullough and Larson, 1999, pp. 133-134).

In the most recent review under study, Koenig, McCullough, and Larson (2001) list 35 studies, which utilised a measure of denomination as a predictor of depression or depressive symptomology. The review reports that certain groups (i.e., Jews, and those not affiliated with a religion) are at greater risk from suffering from depression or depressive symptomology; with the authors declaring, “These findings have been replicated across a number of large, well-designed studies” (p. 135). Similarly, earlier literature reviews (i.e., McCullough and Larson, 1999), which examined the association between religion and depression, came to similar conclusions, namely that, a relationship exists. McCullough and Larson (1999) have claimed that epidemiological empirical studies conducted over the past century tend to demonstrate a trend in results; individuals of particular religious groups have an elevated risk of suffering from certain mental illnesses, such as depression.

The empirical literatures, which compare religious groups, report that certain denominations are more prone to experience depressive symptomology. For instance, Jews (e.g., Levav, Kohn, Golding, and Weissman, 1997; Yeung and Greenwald, 1992) and those people who are not affiliated with any religion (e.g., Brown and Gary, 1987; Ellison, 1995) are more at risk from depression. The results in relation to Christian denominations, such as Catholicism are equivocal, some positive (e.g., Ross, 1990), some negative (e.g., Miller, Warner, Wickramaratne, and Weissman, 1997), and some showing no association (e.g., Levav, Kohn, Golding, and Weissman, 1997). Studies looking at those from Protestant denominations report that individuals involved in the non-mainline Protestant denominations are at an elevated risk of suffering from psychiatric disorders, such as depression (e.g., Meador, Koenig, Turnbull, Blazer, George, and Hughes, 1992). Therefore, it would seem conclusive that those individuals specifically those from a Jewish tradition have a greater likelihood of suffering from depressive symptomology. However, research examining this hypothesis is typically fraught with conceptual and methodological limitations.

As previously stated, Koenig, McCullough, and Larson (2001) conducted an extensive review on the literature investigating the link between religion and health. These reviewers listed all the studies they believed to be relevant to the area and rated them on a scale, ranging from one to ten. One signified a poor study and ten signified an excellent study. In relation to religious denomination and depression, thirty-five studies were listed as belonging to this domain and granted the appropriate rating, according to the reviewers. Out of the thirty-five studies, six (17%) of them received a rating ranging from one to four, indicating that the studies were weak in their design.

Weak Publications (Rating 1-4)

From the six studies considered by Koenig, McCullough, and Larson (2001) to be of a weak design, it was observed that over half of them reported no association between religious denomination and depression (Alvarado, Templer, Bresler, and Thomas-Dobson, 1995; Bazzoui, 1970; Bishop, Larson, and Wilson, 1987; Farakhan, Lubin, and O'Connor, 1984) and under half reported that denomination was positively associated with depression (Flics and Herron, 1991; Payne, Kravitz, Notman, and Anderson, 1976).

The research that reported positive associations between denomination and depression included a study commenting that Jews from a psychiatric population suffered from elevated rates of depressive symptomology (Flics and Herron, 1991). The study conducted by Flics and Herron (1991) consisted of 152 patients drawn from three psychiatric units. The study was designed to assess the contributions of an activity-withdrawal rating (Venables, 1957), a DSM-III diagnoses (American Psychiatric Association, 1980), and the demographic variables of age, sex, race, and religion in predicting pre-morbid competence as measured by the Strauss-Carpenter Prognostic Scale (Strauss and Carpenter, 1974). Results from the study showed that the largest religious affiliation represented in the sample was Catholic, followed by Jewish, and then Protestant; 8% of the sample represented a variety of other religions or professed no religious affiliation. When the religious affiliation was broken down by diagnostic cluster, Protestants in the sample contained the greatest percentage of schizophrenics. The greatest percentage of affective illnesses was found among Jews. Catholics manifested the greatest frequency of disorders which included personality and adjustment disorders and brief and reactive psychotic disorders. No further analyses

were performed specifically on the religious denominations, so, it is not feasible to infer any claims that religious denomination is primarily responsible for the consequence of mental ill health.

Payne, Kravitz, Notman, and Anderson (1976) presented the second study, which showed a positive association between the two variables. The researchers reported that Catholic women suffered greater depression after abortion than women from other denominations. The researchers studied the psychological outcome of abortion in 61 single (M= 23.8 years) and 48 married (M= 31.8 years) female patients. Data from a multivariate analysis suggest that women most vulnerable to conflict are those who are single and nulliparous, those with previous history of serious emotional problems, conflictual relationship with lovers, past negative relationships to mothers, strong ambivalence to abortion, or negative religious or cultural attitudes towards abortion. The authors claim that Catholic women are more likely to have depression following abortion than other denominations ($p < .05$).

Over half of the studies within this section either reported that there were no association between denomination and depression or the studies did not perform statistical analyses on the variables to generate a potential association. One study performed by Bazzoui (1970) examined affective disorders of 98 Iraqi patients. Patients denominations were recorded (90 Muslims, 6 Christians, and 2 Jewish), but no statistical analyses were performed. Bishop, Larson, and Wilson (1987) performed a case-control, cross-sectional study on 64 patients with major affective disorders, compared with 109 unmatched controls with psychiatric disorders. Results showed that denomination was not related to depression.

Similarly, Alvarado, Templer, Bresler, and Thomas-Dobson (1995) assessed the relationship of religious variables to death depression and death anxiety. The sample consisted of 200 respondents obtained from universities, employees and spouses from a hospital in the California area, and managers and sales representatives from a video-distributing company in Los Angeles. Results showed that denomination was not associated to death depression.

It would appear by examining the publications within this subsection that the relationship between religious denomination and depression is one that is ambiguous; with some research pointing to the existence of relationship within certain populations, whilst other research concludes that there is no relationship. When looking at the evidence from this subsection on denomination and depression it would appear that it is not possible to draw a conclusion regarding the relationship between the two variables. Subsequently, subjecting the studies to A-Levels-of-Evidence systematic review, the research would be placed in the C category and regarded as inconclusive as the studies utilise cross-sectional designs with one exception.

One publication within this subsection reported on a longitudinal methodology. Farakhan, Lubin, and O'Connor (1984) conducted a 3-month prospective cohort study assessing the life satisfaction and depression among retired Black persons. The sample consisted of 23 females and 7 males (aged 52-97 years) black retired persons. There was no statistical analysis made between denomination and depression. Therefore, it is questionable why Koenig, McCullough, and Larson (2001) cited this

particular publication as showing a relationship between religious denomination and depression.

Although, the research conducted by Farakhan, Lubin, and O'Connor (1984) was of a longitudinal methodology; the researchers did not utilise statistical analysis or controls. The typical characteristic of the studies in the section is that they omit the employment of controls, except for the study conducted by Alvarado, Templer, Bresler, and Thomas-Dobson (1995). This study, however, is cross-sectional, and does not proceed to be evaluated under the systematic review process. Consequently, all the studies cited by Koenig, McCullough, and Larson (2001) within this subsection appear to be of inadequate quality to be reviewed under the systematic process of A-Levels-of-Evidence approach.

Average Publications (Rating 5-6)

From the review, the authors cited eleven studies (31%), which examined the association between religious denomination and depression. These studies received a rating from Koenig, McCullough, and Larson (2001) of five and six, thus indicating that the research was average in its overall design.

The findings from the eleven studies (45%) report that religious denomination is not associated with depression. This finding was revealed in a study among a sample of 309 members of a religiously liberal (N= 51) and evangelical (N= 77) Protestant churches, a Jewish temple (N= 39), two Unitarian churches (N= 45) and Catholics (N= 97), utilising the Beck Depression Inventory (Genia and Shaw, 1991) and among 86 university students of psychology or religious studies (Plante and Manuel, 1992). Sherkat and Reed (1992) found no association between denomination and depression

among 156 bereaved individuals and Spendlove, West, and Stanish (1984) revealed no differences in scores on depression as measured by the Beck Depression Inventory (Beck and Beck, 1972) between 143 Mormon and 36 non-Mormon white, married women with children aged 14 years or younger. Similarly, Sorenson, Grindstaff, and Turner (1995) found no association between denomination and depression among adolescent unmarried mothers. Finally, Williams and Hunt (1997) reported no difference in depression scores between Muslims and individuals from other denominations.

The studies reporting a positive relationship involved a study examining the relationship between group religious affiliation and mental disorder (Buckalew, 1978). Buckalew (1978) assessed 1,323 institutionalised adult subjects belonging to six major denominations. Chi-square analysis failed to show any differences overall, in functional disorders across religious denominations. However, there were differences in individual disorders in which neuroses and affective disorders were more prevalent among Pentecostals ($p < .01$). Cooklin, Ravindran, and Carney (1983) found differences between Jewish ($N = 64$) psychiatric patients over a 2-year period and non-Jewish patients ($N = 722$) in relation to affective psychoses and affective disorders; symptomatology for these disorders were more prevalent in Jewish patients ($p < .01$). Similarly, Malzberg (1973) reported in a sample of over 40,000 psychiatric patients in New York, depression and manic depression were more prevalent in native-born Jews than non-Jews ($p < .01$, $p < .05$, respectively) and more prevalent in foreign-born Jews compared to non-Jews ($p < .001$, $p < .05$). Jensen, Jensen, and Wiederhold (1993) revealed in a cross-sectional study of 3,835 university students that Mormon women reported less depression compared to non-Mormon women

($p < .001$). Finally, Lubin, Zuckerman, Breytspraak, Bull, Gumbhir, and Rinck (1998) reported that individuals affiliated with a religion were more likely to report less negative affect than individuals who were not affiliated with any religion.

It would seem that there is some evidence to support the proposition that certain denominations have an elevated chance of being associated with higher levels of depression, particularly non-mainline Protestants and those from Jewish affiliations. Although a good proportion of the studies employed controls in their analysis, the vast majority of the studies are of a cross-sectional nature making them invalid under A-Levels-of-Evidence approach, and consequently it cannot be concluded whether mental ill health is a cause or consequence of religious denomination. However, Sorenson, Grindstaff, and Turner (1995) offer an exception to this section as they employed a prospective cohort design and controls such as family composition, socio-economic status, marital status, and social support scale. In addition, they meet all the criteria for evaluation by the approach.

A two-year study conducted by Sorenson, Grindstaff, and Turner (1995) examined religious involvement among unmarried adolescent mothers. The study consisted of teenage mothers ($N = 261$), who replied to a number of questions including religious denomination and scores on the Center for Epidemiologic Studies Depression Scale (Radloff, 1977). Results revealed that respondents who reported that they belonged to 'Other Faith', had the highest depression scores ($M = 18.94$), followed by Catholics ($M = 17.09$), then the group reporting they had no particular faith ($M = 15.52$), and finally by Protestants ($M = 15.18$). When the group was stratified by marital status, it was revealed that the highest depressive symptomology score in the unmarried

category was reported by Other Faith (M= 23.20), followed by Catholics (M= 19.08). However, in the married category, Catholics reported the lowest depressive symptomology scores (M= 8.35). Further analysis revealed that for the unmarried respondents, the average depression scores of women in the more conservative Catholic and Other categories are significantly higher than the scores of Protestants and those with no particular faith ($t= 2.69, p < .01$).

The evidence from the publication performed by Sorenson, Grindstaff, and Turner (1995) would suppose that certain persons affiliated with particular denominations are at an elevated risk of suffering from distress. In this instance, the research would assume that Catholic unmarried teenage mothers and members of other denominations, primarily Pentecostal, Jehovah's Witnesses, or other Christian denominations, suffer more distress than the less conservative denominations. However, the research did not provide any statistical analysis for confounding variables, which may have an indirect effect on the relationship. For instance, it could be possible that demographic characteristics and social variables play an intervening role between the relationship; a phenomena that has been inherent in research examining the religion and health association (McCullough and Larson, 1999). In terms of the systematic review, although the research performed was of a longitudinal methodology, the failure to incorporate confounding variables in the analysis, represents a major flaw that prohibits evaluation of the evidence to maintain whether a justifiable relationship exists or not. Therefore, the study is rated a C, which denotes that the evidence provided is inconclusive.

Good Publications (Rating 7-8)

Koenig, McCullough, and Larson (2001) consider that 42% of the studies investigating the link between religious denomination and depression are good in their overall design. The authors rate these thirteen studies as seven or eight on their scale. Findings from the majority of the studies report that religious denomination is positively associated with depression for Jews. For instance, Fernando (1975, 1978) reported between a sample of 46 Jewish and 71 Protestant depressives and a total of 117 Jewish and Protestant non-psychiatric controls that religiousness was associated with depression among Jews but not among Protestants. Hertsgaard and Light (1984) reported that Catholic women scored highest on depression, Lutheran women second and women from other faiths lowest. Levav, Kohn, Golding, and Weissman (1997) in a sample consisting of 5,772 respondents surveyed in the Epidemiologic Catchment Area study in Los Angeles and New Haven reported that Jews had a higher lifetime prevalence of major depression than Catholics ($p < .05$). In addition, stratifying the groups by gender revealed that Jewish males had significantly higher rates of depression than Catholic, Protestant, and all non-Jews combined.

Similarly, Ross (1990) and Yeung and Greenwald (1992) reported differences in major depression among Jews compared to Catholics and Protestants. For instance, the study conducted by Ross (1990) considered by McCullough and Larson (1999) to be of high epidemiological value, consisted of a probability sample of 401 Illinois residents, aged between 18 and 23. Respondents were asked to comment on a variety of questions assessing their psychological distress, measured by symptoms of depression and anxiety, religious preference (Protestant, Catholic, Jewish, Other, or no religion), and a number of socio-demographic characteristics (age, education,

marital status, family income, gender, and race). Results reported from the analysis revealed that Protestants and those reporting no religious affiliation had the lowest distress score, followed by Catholics ($p < .05$), Jews ($p < .05$), and others ($p < .01$), who had the highest distress levels (uncontrolled analysis). In an controlled analysis, Yeung and Greenwald (1992) revealed from a sample of individuals surveyed in the National Institute of Mental Health Epidemiologic Catchment Area program that Jewish individuals were more likely to have higher rates of major depression than Catholics (CI 1.2-2.4) or Protestants (CI 1.0-2.1), with adjustments for age, gender, race, and socio-economic status.

Most research within this subsection generally employs community samples, which consist mainly of white participants and the sample is overwhelmingly female. However, Brown and Gary (1994) examined the association between religious denomination (affiliates vs. non-affiliates) and depression in a sample of 537 African-American men and revealed that there were significant mean differences for scores on the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) scale between males affiliated ($M= 11.76$) and males not affiliated ($M= 16.95$), at the $p < .05$ significance level. In addition, when selected socio-demographic characteristics were held constant, religious denomination remained significant.

Another feature of the majority of research examined within this subsection is that they typically assess Judeo-Christian denominations, without splitting the denominations into their specific branches. In being more denominationally specific, Meador, Koenig, Hughes, Blazer, Turnbull, and George (1992) found that a higher prevalence of depression was reported among Pentecostals than respondents from

other denominations. The researchers utilised data from the Duke Epidemiologic Catchment area survey to examine the relationship between religious affiliation and major depression among 2,850 adults in the community. Religious affiliations were categorised into six groups; mainline Protestants, conservative Protestants, Pentecostals, Catholic, other religions, and no affiliations. The 6-month prevalence of major depression among Pentecostals was 5.4% compared with 1.7% for the entire sample. Even after psychosocial factors such as age, gender, race, socio-economic status, negative life events, and social support were controlled, the likelihood of major depression among Pentecostals was 3-times greater among persons from other denominations.

In addition, one of the twelve studies stated that religious denomination was unrelated to depression (Braam, Beekman, Knipscheer, Deeg, van den Eeden, and van Tilburg, 1998). This finding was in direct opposition to the findings of Meador, Koenig, Hughes, Blazer, Turnbull, and George (1992). Few empirical publications within this subsection controlled for a vast number of intervening variables. When covariates were controlled for, the strength of the association is typically reduced to nonsignificance. In addition, the majority of research to this point, looking at denominations, typically clusters the Protestant denominations together, although Protestantism is one denomination that has many branches, which run on a spectrum from conservative to liberal. Consequently, one piece of research performed by Braam, Beekman, Knipscheer, Deeg, van den Eeden, and van Tilburg (1998) splintered the branches of Protestantism into three separate categories, whilst examining the role of a plethora of covariates.

The researchers reported on a sample of 3,020 elderly community dwelling adults, between the ages of 55 and 85, and reported that liberal Protestants had higher depression scores than respondents from other denominations. The Protestant section of the sample was split into three separate categories: Reformed Calvinist, Orthodox Calvinist, and Other Protestant denominations (Lutherans, Baptists, Mennonites, Old Catholic Church, Pentecostal Church, Free Evangelical Communities, and Full Evangelical Communities). The remaining categories were Roman Catholic and no denomination. The respondents indicated their level of depressive symptomology, utilising the Center of Epidemiological Depression Scale (Radloff, 1977). In addition, information on a number of background variables and covariates were obtained (i.e., social integration, sense of mastery, self-esteem, socio-demographics, and functional limitations). With respect to the prevalence of depressive symptoms, Reformed Calvinists have the lowest mean depression score (6.9), followed by the Dutch Reformed (7.4), and Roman Catholics (7.5). These differences, as compared to nonchurch members, remain significant after controlling for the effects of demographic variables (statistics not shown). The category, Other Protestants had the highest mean depression score (10.7), although, the difference with the reference group was no longer significant after controlling for demographic variables (marital status, age, and degree of urbanisation).

Further analysis by Braam, Beekman, Knipscheer, Deeg, van den Eeden, and van Tilburg (1998) explored the associations between denomination and explanatory variables, which revealed a number of observations. Firstly, it was shown that both Calvinists and Roman Catholics had larger social networks and received more emotional support. Self-esteem scores were higher among Roman Catholics ($\beta = .06$,

$p < .05$), whereas they were lowest among Calvinists ($\beta = -.06, p < .05$). In examining the association between the variables, the researchers revealed that significant negative associations between the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) scores and both Calvinist ($\beta = -.06, p < .05$) and Roman Catholic ($\beta = -.06, p < .05$) background were found, after controlling for effects by demographics and functional limitations. For Calvinists, controlling for social integration variables decreased the strength of the association with the depression scale score, but this did not happen for Catholics. In the final model, the inclusion of church attendance to the equation decreased the strength of the association for both denominations, to nonsignificance. Therefore, it can be postulated that church attendance exerts an intervening effect in the association between religious denomination and the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) score.

Evidence provided by the publications within this subsection would propose that certain denominations are at an increased risk from depressive symptomology, in particular those from a Jewish tradition. However, the publications which present such conclusions come from research that is methodologically weak, in that they employ cross-sectional designs and generally provide such conclusions without the employment of controls, which may explain the observed relationship. Regarding these flaws, the publications in the subsection would be rated in the category C, which represents evidence that is flawed and inconclusive.

Although the majority of the research within this subsection was of a cross-sectional methodology, few publications represented results from longitudinal studies. For instance, the first study performed by Williams, Larson, Buckler, Heckmann, and Pyle

(1991) examined the association between religious denomination and psychological distress in a 2-year community study of 720 adults. The main predictor variable, religious denomination was measured by respondents indicating if they were affiliated with a church or religious group and the outcome variables, psychological distress, was measured by Gurin, Veroff, and Feld (1960) symptom checklist scale, which indicates the presence of moods of anxiety and depression. In addition, information was obtained on socio-demographic characteristics and stressful life events (negative life events and physical health limitations). Individuals who indicated that they were affiliated with a religion did not differ significantly from those individuals who reported that they were not affiliated with a religion on scores of depression, as measured by the symptom checklist scale, at two time points.

Miller, Warner, Wickramaratne, and Weissman (1997) conducted the second longitudinal publication within this subsection. The research was a 10-year prospective cohort study investigating the relationship between religiosity and depression. The sample consisted of 60 mothers and 158 of their offspring, who provided information on a variety of questions, including their religious denomination at two time points and depression. The study was performed to test several hypotheses. Firstly, was religiosity protective against the prevalence of depression in mothers. Secondly, was offspring religiosity protective against the prevalence of depression. Thirdly, was maternal religiosity protective against prevalence of depression in offspring. Finally, was mother-offspring concordance of religiosity was protective against prevalence of depression in offspring. In the study, the offspring of depressed and nondepressed parents were followed for ten years. From the analysis it was revealed that at time 1, there were no significant differences between Catholics

and Protestants in relations to time 1 maternal major depressive disorder. However, at time 10 (follow-up), compared with a mother who was Protestant, a mother who was Catholic was 79% less likely to have major depressive disorder at time 10 (OR= 0.21, CI 0.07-0.7, Wald's $\chi^2= 7.0$, $p < .01$). This finding remained consistent with adjustments for time 1 depression and marital status, ethnicity, and social functioning (OR= 0.14, CI 0.04-0.5, Wald's $\chi^2= 8.2$, $p < .01$). There was no significant associations to support the hypothesis that offspring religiosity was protective of offspring major depressive disorder. It was also revealed that maternal religiosity and offspring depression was not statistically significant as was mother and offspring concordance. The findings from the analysis reveal that among mothers, religious importance and Catholicism are protective against prevalence of major depressive disorder.

It can be concluded that the majority of the studies report positive associations between denomination and depression. However, the findings were ambiguous, with no specific denomination except Judaism providing consistent results. The studies that reported positive associations between denomination and depression performed multiple controls with the exception of Fernando (1975, 1978). However, with a closer examination of the literature it can be observed that the majority of the studies controlled for low-level socio-demographic variables, such as gender, age, marital status, education, and race (Brown and Gary, 1994; Jones-Webb and Snowdon, 1993; Levav, Kohn, Golding, and Weissman, 1997; Ross, 1990; Yeung and Greenwald, 1992). Nevertheless, few of the studies showing the positive association did control for covariates in a high-order analysis. In relation to A-Levels-of-Evidence approach, the majority of the studies would be placed in the C category signalling that the

evidence from these studies were inconclusive. The majority of the studies were cross-sectional studies, and of the two that incorporated longitudinal designs one (Williams, Larson, Buckler, Heckmann, and Pyle, 1991) reanalysed data, which was utilised earlier on the same cohort and therefore receives a rating of C (inconclusive). Consequently, that leaves one study within the section that could provide adequate evidence to generate a conclusion on the effects of denomination on depression. The study that is considered to be of adequate standard to be included for evaluation by the systematic review was performed by Miller, Warner, Wickramaratne, and Weissman (1997). However, due to the study not incorporating protective factors, such as social support, the model is evaluated under the mediated model and not the independent model. Therefore, the publication receives a rating of A (positive) signifying that the evidence presented by the publication was conclusive and did not contain any of the flaws, and that Catholicism was protective against depression in mothers.

Excellent Publications (Rating 9-10)

Koenig, McCullough, and Larson (2001) cited few studies from the review (16%) that assessed religious denomination and depression, which received a rating of very good (9) or excellent (10). From these studies, one performed a cross-sectional design (Koenig, George, Meador, Blazer, and Dyck, 1994) with the remaining incorporating prospective cohort designs (Idler and Kasl, 1992; Kennedy, Kelman, Thomas, and Chen, 1996; Koenig, Cohen, Blazer, Pieper, Meador, Shelp, Goli, and DiPasquale, 1992). Findings from these studies revealed a positive relationship between denomination and depression; however, whether or not this is warranted needs examination.

Koenig, George, Meador, Blazer, and Dyck (1994) performed the first study, which examined depression in non-mainline Protestants, wherein the prevalence rates of depression between Pentecostals to the rates of depression among members of other Christian faiths were compared. The study was based on data obtained in waves I and II of the National Institute of Mental Health's Epidemiologic Catchment Area survey in North Carolina. From this survey, 2,679 respondents were utilised. The sample was split into two ways, "baby boomers" and middle-aged, and older adults. All participants were administered the Diagnostic Interview Schedule (Robins, Helzer, Croughan, and Ratcliffe, 1981), data was also collected on religious affiliation and attendance at religious services. In addition, a wide range of socio-demographic and health information were collected. Socio-demographic data included age, sex, race, marital status, and socio-economic status. Physical health status was assessed by the presence or absence of one or more chronic illnesses.

Results demonstrated that compared with members of other conservative denominations or mainline Christian denominations, the 6-month and life-time risk of depressive disorders for the Pentecostal "baby boomers" was considerably higher compared to "baby boomers" from other conservative denominations or from mainline Christian denominations. These differences in prevalence held even after controlling for sex, race, physical health, and socio-economic status. Additionally, Pentecostals had the highest rate of any recent DSM-III psychiatric disorder (27.3%) as well as lifetime disorder (42.8%, $p < .01$). This remained significant after controlling for covariates. Pentecostalism was not associated with differential prevalence rates of depression in the older age group.

One of the criticisms concerning the research performed on religion and health is that the research generally employs cross-sectional designs. Prohibiting any potential causal relationship being inferred from the analyses. Therefore, the previous study mentioned within this subsection would receive a rating of C, denoting that the evidence presented by the publication was inconclusive, under A-Levels-of-Evidence approach. Reviewers have provided opposing conclusions to the results reported by the investigators utilising longitudinal designs. For instance, McCullough and Larson (1999) argue that longitudinal studies disagree with findings from the former cross-sectional type studies. These researchers argue against the general agreement that being Jewish is associated with greater psychological distress and depressive symptomology.

Nevertheless, the majority of the studies within this section, incorporating prospective cohort designs, revealed mixed findings. For instance, Kennedy, Kelman, Thomas, and Chen (1996) conducted a study that included a longitudinal analysis to examine the relation of religious preference to late life mental illness. Kennedy, Kelman, Thomas, and Chen (1996) interviewed 1,855 older community residents (65+). The study specifically assessed the differences in the prevalence and course of depressive symptoms among residents who expressed Jewish (n= 711), Catholic (n= 880), or “other religious” preference (n= 264). During baseline in-person interviews, respondents were asked to indicate their religious preference as Catholic, Jewish, Protestant, none, or other and were asked to complete the Center for Epidemiologic Studies Depression scale (Radloff, 1977). In addition, respondents provided information on health, chronic illness, physical and cognitive functioning, utilisation

of and attitudes towards health care, interactions with family, friends, and social services agencies, and financial resources.

The overall statistical analyses, at baseline, showed that individuals who expressed a Jewish religious preference were associated with a two-fold rise in depressive symptoms compared with Catholics. It was reported that Jews were more likely to score higher on the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) than Catholics or other faiths (20.7% vs. 9.5% and 12.3%). This finding remained significant when the effects of age, gender, immigration, health, disability, and social support were taken into account. It was observed that using logistic regression to control for other covariates, Jews were 75% more likely than persons of other affiliations to have depression (OR 1.75, CI 1.51-2.01, $p < .001$). Additional variables found to be statistically significant in the model were, fair-to-poor health ($p < .001$), problems with activities of daily living ($p < .001$), living alone ($p < .001$), non-attendance of religious services ($p < .001$), female gender ($p < .01$), receipt of both formal and informal social support services ($p < .05$), education equal to or more than 9 years ($p < .01$), and having 2 or more cardiovascular conditions ($p < .05$).

Twenty-four months after the baseline interviews, 85% of the original sample completed a second administration of the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) and responded to questions regarding changes in problems with activities of daily living, social support, and whether their health improved, stayed the same, or declined. Results report that significantly more Jewish respondents than Catholic respondents or respondents from the other group experienced an emergence of depression. In addition, Jewish respondents within the

sample made up a greater percentage of persons in whom depression persisted. However, these differences were not significant compared to Catholic respondents and respondents who expressed themselves as belonging to the other group.

Kennedy, Kelman, Thomas, and Chen (1996) not only reported the over representation of Jewish adults at baseline, they also remarked on the amount of Jews who experienced depression at the 24-month follow-up. However, when the researchers performed stepwise and canonical discriminant analyses, they found that the factors which were most likely to contribute to the emergence of depression were problems with daily physical capabilities ($F= 99.83, p< .001$), the respondents' health (if it gradually became poorer over the forthcoming year; $F= 46.01, p< .001$), and if the respondents rated their health as fair-to-poor at baseline ($F= 9.70, p< .001$). Alternatively, the variable "being Jewish" only accounted for less than 1% of the variance as a characteristic distinguishing respondents who were never depressed from those in whom depression emerged ($F= 6.51, p< .01$).

The conclusions rendered by Kennedy, Kelman, Thomas, and Chen (1996) are inconsistent with the results reported by the sociologists, Idler and Kasl (1992). Idler and Kasl (1992) conducted a longitudinal study involving 2,812 older adults (65+) over a three-year period from New Haven. It was reported that being Jewish at baseline actually lead to a lower risk of developing depression. Depression was assessed through the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) and the results remained constant even after controlling for a variety of socio-demographic covariates, such as perceived income adequacy, race, education, sex, and age, as well as for controlling for the baseline levels of depression. Idler and Kasl

(1992) assessed the impact between the relationship of religious denomination and depression for a number of intervening variables. These intervening variables contained a set of health practices (e.g., smoking, alcohol consumption, exercise, and body mass index), as well as three indicators of social networks (the number of people seen frequently, the number of close friends and family members, and a dummy variable of being married), and an indication for an attitude of optimism. Even after controlling for these psychosocial variables, results remained consistent.

The study conducted by Idler and Kasl (1992) is considered by the reviewers (Koenig, McCullough, and Larson (2001) as being one of the pre-eminent studies within the area. The results from the final analysis, in relation to depression, revealed that Jews were considerably less likely than members of the other religious denominations to become depressed between 1982-1985 ($p < .05$). Additional variables that predicted depression over the period were functional disability ($p < .001$), change in disability from 1982-1985 ($p < .001$), and being married in 1982 ($p < .001$).

The researchers concluded that the findings from the study suggest that religion does play a profound affect on the health of the elderly. They believe this is represented within the design of the study. Idler and Kasl (1992) utilised a large and representative sample, used multiple measures of both religion and health status, conducted a prospective study design, utilised conservative techniques for the analysis of the data from stratified samples, and used a range of controls. This design suggests that the observed affects of religion were not due to selection or to the inadequate control for other confounding factors. However, methodological problems in reference to the measures utilised requires discussion. For instance, measures of

religious denomination suffer from certain conceptual flaws (i.e., does religious denomination capture the concept of religiousness), which threatens its validity (Levin and Schiller, 1987) and operationalisation of the depression construct, as measured through the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) encounters problems with validity for certain populations.

One study representing the effects of belonging to a religious denomination outside Judaism, in relation to depression was conducted by Koenig, Cohen, Blazer, Pieper, Meador, Shelp, Goli, and DiPasquale (1992). The researchers examined religious coping and depression among 850 medically ill men aged 65 years and over, without psychiatric diagnoses. Measures employed in the study consisted of the self-rated 30-item Geriatric Depression Scale (Yesavage, Brink, Rose, Lum, Huang, Adey, and Leirer, 1983) and the observer-rated Hamilton Rating Scale for Depression (Hamilton, 1967), and three items assessing religious coping. In addition, data was collected on demographic characteristics including, race, age, education, prior occupation, retirement status, current living situation, marital status, yearly income, religious denomination, and social support. A 3-item index measured social support that explored size of social networks, frequency of interaction, and perceived adequacy of support. Finally, the data was obtained examining the physical health of the participants, including medical diagnoses (i.e., chronic illnesses), functional status (i.e., physical and instrumental activities of daily living), and cognitive status.

Religious denomination was categorised into nine general groups: Liberal Protestant (8%), moderate Protestant (12.5%), conservative Protestant (40.8%), black Protestant

(22.2%), Fundamentalist/evangelical (5.1%), Protestant (unspecified= 3.3%), Catholic (2.6%), non-traditional Christian (3.2%), and no religious preference (2.4%).

Results from bivariate correlations showed that, with regard to the Geriatric Depression Scale (Yesavage, Brink, Rose, Lum, Huang, Adey, and Leirer, 1983), being conservative Protestant was positively associated with depression ($r = .11$, $p < .001$), whilst being black Protestant was negatively associated with depression ($r = -.10$, $p < .01$). In relation to the Hamilton Depression Scale (Hamilton, 1967), no religious preference is positively associated with depression score ($r = .14$, $p < .01$), but moderate Protestants are negatively related to scores on depression ($r = -.12$, $p < .05$). It can be observed that these values represented a weak association between the two variables. The variable that proved to have the strongest association with the Geriatric Depression Scale (Yesavage, Brink, Rose, Lum, Huang, Adey, and Leirer, 1983) and the Hamilton Depression Scale (Hamilton, 1967), was history of psychiatric problems, $r = .30$ and $r = .27$ ($p < .001$), respectively. Hence, it can be suggested that if an individual has suffered from some form of mental ill health in the past then they are more likely to be suffering from depressive symptoms subsequently.

Further analysis using stepwise regression procedures created two models for depression. The first model (using Geriatric Depression Scale scores as the dependent variable: Yesavage, Brink, Rose, Lum, Huang, Adey, and Leirer, 1983) incorporated age, race, social support, history of psychiatric problems, family history of psychiatric problems, alcohol use, cognitive status, functional status, and six medical diagnoses into the final model. These variables accounted for 26% of the variance in depression

scores. However, when religious denomination was entered into the model it did not contribute significantly to the prediction of depression scores. The second model, which utilised the Hamilton Depression Scale (Hamilton, 1967) scores as the dependent variable, incorporated social support, history of psychiatric problems, alcohol use, and functional status, was able to explain 16% of the variance in depression. Conversely, when religious denomination was added, it contributed significantly to the model, with moderate Protestant, Catholic, and non-traditional Christian emerging as significant predictors. However, this improvement in significance was minor, $F(8,310) = 2.7, p < .01$. The model showed that these denominations were negatively related to scores on depression. In other words, belonging to a moderate Protestant tradition, Catholicism, or non-traditional Christian appeared to exert protective effects against depressive symptoms.

The evidence provided by the publications cited within this subsection would suggest that certain religious denominations are associated with being protected against depressive symptomology, whilst other religious denominations are associated with an increased risk of suffering from depressive symptomology. Cross-sectional studies would generally conclude that individuals from a Jewish tradition are associated with being one of the highest risk groups of suffering from depressive symptomology. Examining results from methodologically sound publications showed that Judaism was generally not associated with depression. However, one study revealed a weak association (Kennedy, Kelman, Thomas, and Chen, 1996). This observed association could be explained by some third variable not controlled for. In terms of the systematic review, the publications presented within this subsection would be rated and evaluated by A-Levels-of-Evidence approach. They do not include the

methodological flaws that prohibit inclusion to the review. Therefore, the publications can be rated and the strength of the evidence for the hypothesis, that religious denomination is protective against depression can be judged.

Examining the hypothesis, which states that religious denomination is associated with lower levels of depressive symptomology, can be examined by a few of the studies. The first study, performed by Kennedy, Kelman, Thomas, and Chen (1996), reported that Jews were more likely to suffer from depression than respondents from other denomination and this finding remained consistent after controlling for a variety of covariates. In relation to the mediated model, the publication would receive a rating of A (negative), denoting that the evidence provided by the study was conclusive. The methodology of the study was judged to be sufficiently sound to support conclusions about their assertions. The study's conclusion was negative. In relation to the independent model, the publication would also receive a rating of A (negative) as it did not contain any of the flaws that would cloud interpretation. It can be postulated that because the relationship between Judaism and depression remains significant in both models, the relationship cannot be completely accounted for by known risk factors.

The second study, performed by Idler and Kasl (1992), presents findings in direct opposition to those presented in the previous study. Idler and Kasl (1992) reported that being Jewish actually lower the risk of suffering from depression. In terms of assessing the evidence by A-Levels-of-Evidence approach, assessing the publication under the mediated model, it can be reported that the publication would be placed in the category of A (positive). This signals that the methodology of the study does not

contain flaws that may cloud interpretation; however, in relation to the hypothesis that Jews are associated with depression the study is positive. Similarly, in relation to the independent model, the category, which the study would be placed, is A (positive), denoting that the methodology is conclusive, and supports the hypothesis positively.

The last study within this subsection assessed the relationship between religious denomination outside Judaism and depression. The study conducted by Koenig, Cohen, Blazer, Pieper, Meador, Shelp, Goli, and DiPasquale (1992) examined the relationship between Christian denominations and those not affiliated with religion alongside depression. The results from the analyses were mixed, with religious denomination not being associated with depression when depression was measured by the Geriatric Depression Scale (Yesavage, Brink, Rose, Lum, Huang, Adey, and Leirer, 1983). However, when the Hamilton Depression Scale (Hamilton, 1967) was utilised it was found that those belonging to a moderate Protestant tradition, Catholicism, or non-traditional Christian appeared to exert protective effects against depression. In terms of the hypothesis, religious denomination is associated with lower depression, the study provides partial support. In relation to the mediated model, the study would receive a rating of A (positive) and in terms of the independent model the study would receive a rating of A (positive). This demonstrates that the methodology of the study meets minimal methodology standards which provided conclusive evidence that religious denomination exerts protective effects on depressive symptomology. Looking at the discrepancies between the findings related to both measures, it can be suggested that because they are measuring different perceptions of depression, could explain the inconsistent findings.

For instance, the Geriatric Depression Scale (Yesavage, Rink, Rose, Lum, Huang, Adey, and Leirer, 1983) is a self-report measure of depressive symptomology, whilst the Hamilton Depression Scale is an observer-rated depression scale.

When examining the ratings provided by evaluation from the systematic review process, on all the subsections, it appears that there are enough high quality studies to infer a conclusion based on the data from the studies. Four studies from this section assessing the relationship between religious denomination and depression are considered to meet the minimal methodological standards to provide an evaluation of the strength of the hypothesis that religious denomination is a protective factor against depression. The four publications were all rated in the A category, with three of them providing positive support for the hypothesis. Therefore, the hypothesis that religious denomination is protective against depression reaches the level of *persuasive evidence*.

Anxiety

Reviews assessing the association between religious denomination and anxiety are ambiguous. Sauna (1969) reported that the literature 'Points out that the religious person may at times show greater anxiety and at times less anxiety' (p. 1206). In a more recent review, Bergin (1983) found that 'Contradictions in results (in research on religion and psychopathology) continued in manifest anxiety' (p. 123). However, this general conclusion is not validated in the review conducted by Koenig, McCullough and Larson (2001). These reviewers claim, 'The preponderance of evidence suggests that religion as a whole... tends to buffer against anxiety' (p. 153).

Koenig, McCullough, and Larson (2001) reported on seventeen studies that utilised a measure of religious denomination and assessed its association with anxiety. These reviewers rated the studies examining religious denomination and anxiety ranging from one to eight, with the lower rated studies indicating research that is weak in design.

Weak Publications (Rating 1-4)

Koenig, McCullough, and Larson (2001) rated five (29%) of the studies in the weak category (1-4). The studies in this group conducted cross-sectional research with the exception of Raphael, Rani, Bale, and Drummond (1996), in which a case-control study was performed. The findings from these studies report mixed findings. For instance, some researchers state that religious denomination is not related to death anxiety (Alvarado, Templer, Bresler, and Thomas-Dobson, 1995; Iammarino, 1975) and general anxiety (Kutner, 1971). However, other researchers report that religious denomination is positively associated with obsessive-compulsive disorder (Raphael, Rani, Bale, and Drummond, 1996).

The studies presenting no association included research examining the relationship of religious variables to death depression and death anxiety. The research conducted by Alvarado, Templer, Bresler, and Thomas-Dobson (1995) was a cross-sectional study consisting of a heterogeneous group of non-clinical subjects. The sample comprised 200 subjects (114 females and 86 males) obtained from four sources: higher education institutions, employees and spouses from a medical centre and a hospital in California. In relation to the variables under interest, the subjects were administered the Death Anxiety Scale (Templer, 1970) and the denomination of the respondent was

recorded. The denominational background of the combined group was 92 (46%) Catholic, 45(23%) Protestant, 10 (5%) Jewish, 5(3%) non-believers, and 48(24%) belonged to 'Other' religions. There were no statistical analyses performed between denomination and death anxiety; therefore, not rendering any conclusion being drawn on these two variables from the research. However, Immarino (1975) assessed the relationship between death anxiety, utilising the Templer's Death Anxiety Scale (Templer, 1970), and demographic variables, including religious denomination. The study consisted of 249 high school students, including 58 Catholics, 8 Jews, 83 Protestants, and 75 Others. Results present no association between denomination and death anxiety in high school students. Similarly, when Kutner (1971) examined fear of pregnancy, results from the research revealed that denomination was not related to fear of pregnancy.

The study showing a positive relationship between the variables, examined the effects of religion and ethnicity on obsessive-compulsive disorders, a common anxiety complaint that affects individuals from diverse cultures as India, Hong Kong, China, as well as the Western countries (Raphael, Rani, Bale, and Drummond, 1996). The study was performed in London, England and consisted of three groups of patients, 49 patients diagnosed with obsessive-compulsive disorder (OCD) and two control groups. The first control group consisted of 51 referrals to the General Adult Psychiatric Services; the second group included 48 referrals from the Specialist Psychotherapy Service. Information was obtained from referral letters, clinical notes, and subsequent correspondence, which possessed data on religion, ethnic group, country of origin, and social class.

Results from the study show that individuals in each of the groups were more likely to report that they belonged to a religious denomination, 75.5% in the obsessive-compulsive disorder group, 77.1% in the psychotherapy group, and 52.9% in the general psychiatry group. In comparing the three groups, chi-square analysis revealed modest differences in the number of individuals identifying themselves with any religion ($\chi^2 = 11.0$, $df = 4$, $p < .05$). The majority of the sample (60.8%) identified themselves as either Catholic or another form of Christian. However, there was no significant difference in the type of religion practised by the three experimental groups ($\chi^2 = 11.4$, $df = 8$, ns). The conclusions presented from this study were indefinite. Although it was revealed those from a religious background belonged to the OCD group or the psychotherapy group the analyses does not employ controls or confounding variables.

The majority of the studies within this section reveal that religious denomination is not associated with anxiety. The study, which stated that there were positive associations between denomination and anxiety fail to control for any potential covariates, therefore, rendering validations surrounding the association's null. In relation to evaluating the studies under the systematic review process, the research within the weak rating (as designated by Koenig, McCullough, and Larson, 2001) would all receive a category of C (inconclusive). The studies are all of a cross-sectional nature, which fails to meet the one of the requirements of the review process. However, the study, which conducted a longitudinal study (Raphael, Rani, Bale, and Drummond, 1996), did not employ controls within the analysis, therefore, rendering the evidence as inconclusive.

Average Publications (Rating 5-6)

The majority of the studies (59%) which examined the association between denomination and anxiety were categorised as average (rating 5-6). According to Koenig, McCullough, and Larson (2001), the ten studies generally report that 'religion as a whole' (p.153) offers a beneficial function against anxiety. However, the authors do not specifically state the effect that religious denomination bears on anxiety. Koenig, McCullough, and Larson (2001) list the studies and state if the findings are positive, negative, mixed, or show no association. However, in closer examination of the studies that utilise religious denomination, as a measurement variable, the majority of them report that denomination is not associated with anxiety (Aday, 1984-1985; Bohannon, 1991; Frenz and Carey, 1989; Minear and Brush, 1980-1981; Plante and Manuel, 1992; Steketee, Quay, and White, 1991). These represent over half of the studies in this subsection. Significant findings from the remaining studies include, death anxiety being higher in homosexual men (with and without aids) who state belonging to the same religious denomination as in childhood (Franks, Templer, Donald, Cappelletty, and Kauffman, 1990-1991). Anxiety being greater for individuals who have had a glossolalia experience, a common characteristic in fundamentalist denominations (Kirkpatrick and Shaver, 1992). Finally, death anxiety being associated with young adults affiliated with a religion (Richardson, Berman, and Piwowarski, 1983). Out of the ten studies listed in this section, only one reported the beneficial effect that religion exerts on anxiety (Young and Daniels, 1980).

Young and Daniel's (1980) study examined 'Born-Again' status as a factor in death anxiety among high school students in the rural south. The study focused on the

relationship of religiosity to the fear of death amongst high school students. The sample consisted of 312 students, 120 males and 192 females, from two rural high schools in Alabama. Among the variables of interest, the authors utilised items on religious affiliation, such as born-again status, as a component of the religiosity construct. Religious affiliation was classified according to the procedures of Peterson and Mauss (1976): liberals (Congregationalists, Northern Methodists, Episcopalians, Friends, Unitarians, and members of the United Church of Christ), moderates (Disciples of Christ, Presbyterians, Southern Methodists, Christians, and members of Community Churches), conservatives (Lutherans, American Baptists, Christian Scientists, and Mormons), and fundamentalists (Southern Baptists and Pentecostals). Students were classified as born-again Christians, Christians but not born again, and not Christians. The dependent variable, death anxiety, was measured using the Templer Death Anxiety Scale (Templer and Dotson, 1970). The anxiety scale consists of fifteen items that measured personal experiences such as death concern, fear of the dying process, and fear of corpses. The scores ranged from '0' (no death anxiety) to 15 (very high death anxiety). In addition, data was collected on race and sex.

A stepwise multiple regression analysis was performed to assess the relationship between the variables. Findings indicated that in general females reported higher levels of death anxiety ($M = 8.22$). The authors reported that sex, religiosity, as measured by the intellectual and ideological scales of the Faulkner and DeJong Religiosity in 5-D scale (Faulkner and DeJong, 1966), religious affiliation, Born-again status, and race explained 14.4% of the variance in scores of death anxiety. In relation to religious affiliation, it was reported that religious affiliation explained 2.6% ($p < .05$) of the variance in scores on the anxiety scale and that subjects with

liberal denominational affiliations tended to exhibit greater death anxiety ($r = -.14$). In addition, it was revealed that individuals that reported they were Born-again Christians exhibited lower levels of death anxiety. Although, religious denomination was statistically significant the variables that accounted for the greatest amount of variance in anxiety scores were ideological scale of the religiosity scale ($R^2_{\text{change}} = .040$, $p < .05$) and sex ($R^2 = .033$, $p < .05$).

The remaining studies within this subsection revealed the lack of impact that religion had on anxiety scores. For instance, studies that showed no association between the two variables, included studies assessing student samples (Aday, 1984-1985; Frenz and Carey, 1989; Minear and Brush, 1980-1981; Plante and Manuel, 1992), one study examining the relationship with bereaved parents (Bohannon, 1991), and one study assessing psychiatric patients (Steketee, Quay, and White, 1991).

The typical characteristic of these publications within this subsection, assessing religious denomination and anxiety, is that the researchers conducted cross-sectional research with the majority of the studies not utilising controls or potential confounders in their analysis. With this in mind, it can be declared that the evidence offered is not of adequate standard to be assessed by A-Levels-of-Evidence approach. Therefore, these publications fall under the C category meaning that the evidence is inconclusive.

Good Publications (Rating 7-8)

The highest rating offered by Koenig, McCullough, and Larson (2001) within the anxiety domain was seven and eight. Few studies, approximately two (12%), received this rating. Similar to the previous category, one of the studies did not report a positive association between religious denomination and anxiety (Williams, Larson,

Buckler, Heckmann, and Pyle, 1991), whereas one of the studies did report a significant positive association with anxiety (Hertsgaard and Light, 1984).

Hertsgaard and Light (1984) examined anxiety in rural women. The study was cross-sectional and consisted of 760 randomly selected females residing on farms in a Midwestern state. The subjects' average age was 44 years and 97% of the sample was married. The subjects responded to the Multiple Affect Adjective Check List (Zuckerman and Lubin, 1965) in which scores on the anxiety, depression, and hostility subscales operated as the dependent variables. The independent variables were obtained through the use of demographic and a biographic questionnaire. In relation to religious denomination and anxiety, it was observed that Catholic women scored highest on anxiety ($M= 6.50$, $SD= 4.33$), Lutheran women second ($M= 6.14$, $SD= 4.30$), and women of other faiths lowest ($M= 4.99$, $SD= 4.11$). However, no controls or potential confounders were employed in the analysis.

Conversely, Williams, Larson, Buckler, Heckmann, and Pyle (1991) conducted a prospective cohort study on 720 adults from the New Haven, Connecticut community. The aim of the study was to enhance the understanding between religiosity and mental health by examining how two measures of religious involvement, religious attendance and religious affiliation, combine with stress to affect psychological distress. The sample of 720 adults were re-interviewed in 1969 from an original random sample of 938 respondents interviewed in 1967. The Gurin, Veroff, and Feld (1960) Symptom Checklist Scale measured psychological distress and religious affiliation was assessed by asking respondents, 'Are you affiliated with any church or religious group' (1= yes, 0= otherwise). Various socio-demographic variables were recorded, such as age,

education, gender, marital status, and race. Ordinary Least Squares regression analyses was utilised to estimate the magnitude and statistical significance of the relationships among religious involvement and psychological distress.

Findings from the ordinal least squares regression analysis revealed at Time 1 (1967) religious affiliation did not predict mental ill health as measured by the Gurin (1960) scale, with the introduction of the confounding variables at Time 2 (1969), religious affiliation remained unrelated.

The first publication within this subsection (Hertsgaard and Light, 1984) performed a cross-sectional design, therefore, not comprising the proper standards for the review. Although, the latter study performed a prospective cohort design and utilised multiple controls in its analysis, the published article is of an insufficient standard to be evaluated under the systematic review process as it obtained its data from interview data collected earlier and reanalysed. The data used within the study was reanalysed, so, complying with the elimination criteria, which postulates exclusion if that data employed had earlier reports on the same cohort.

At first glance, it would seem that research assessing the relationship between religious denomination and anxiety appears to generate a relationship that is mixed. Some of the studies report the beneficial effects of religion; some report the detrimental effects of belonging to a religious denomination, whilst others report no association. However, it must be noted that the majority of the empirical research in this branch of religiosity and mental health utilised cross-sectional designs, does not examine the role of confounding variables or stratification variables. Failure to control

for these factors can lead to a biased estimation of any positive relationship. It is important to place controls in the analysis as Levin and Schiller (1987) argue that without adequate methodological controls the “religious factor” in health might be mistakenly dismissed as merely a proxy measure for a host of additional variables, such as socio-demographic status or personality attributes.

In examining the hypothesis that religious denomination protects against depression, under A-Levels-of-Evidence approach, it can be viewed that within the subsections, none of the research meets the minimal methodological standards to be evaluated. The accumulated research reaches the level of *insufficient evidence* for the hypothesis. From the seventeen studies that examined the relationship between religious denomination and anxiety, none of the publications meet the methodological standards to be rated. Investigating research of this standard under A-Levels-of-Evidence strategy denotes that there is not enough evidence for a conclusion to be offered. This signifies the need for a greater amount of research to be conducted of an adequate standard.

Subjective Well-being

Research examining the relationship between religious denomination and depression, has generally revealed the positive effects of this variable on the negative aspect of mental health. Similarly, research on religious denomination suggests that membership or other ties to particular types of religious communities can shape individual assessments of life quality (Ellison and Gay, 1990). Synonyms for subjective well-being typically include life satisfaction, morale, quality of life, self-esteem, hope, optimism well-being and happiness.

Koenig, McCullough and Larson (2001) performed a systematic review of the literature and uncovered 100 studies that statistically examined the relationship between religion and well-being. The authors stated that of these 100 studies investigating the relationship, seventy-nine reported at least one positive correlation, thirteen found no association, one found a negative association, and the remaining seven reported mixed or complex relationships. The number of studies examining the relationship between religious denomination and well-being is moderate, accounting for 21% of the total studies listed. The researchers state that the empirical studies assessing the effects of religious denomination on subjective well-being typically report a positive association (e.g., Gee and Veevers, 1990; Glik, 1986; Reed, 1991).

Weak Publications (Rating 1-4)

In the review, the authors report on one study (Flaskerud and Uman, 1996) that examined the relationship between denomination and positive mental health, which received a rating of two. This study was assessed as poor in terms of overall design. The study was performed to assess acculturation and its effect on the self-esteem among immigrant Latina women. Under A-Levels-of-Evidence approach, the published article is considered of adequate standard to be evaluated since it possesses the necessary criteria not to be placed in the eliminated category. The study attempts to control for adequate confounders, utilises a prospective cohort design, adequately measures religiosity (denominations), performs statistical analyses, and uses data that is been reported on in an earlier cohort.

Flaskerud and Uman (1996) performed a prospective cohort study among 491 immigrant Latin American women, attending a Public Health Foundation Nutrition Program for Women, Infants and Children. The religious variable was measured by denomination and the sample was divided into Catholics vs. non-Catholics. Results show that religious denomination was not associated with self-esteem. However, the religious variables appeared as background variables rather than host variables.

The fact that the religious variable acted as a background variable impacts on the evaluative quality of the study (under the regulations of A-Levels-of-Evidence strategy) in that there was no controls utilised, hence it is not possible to determine if the strength of the relationship could be explained by some other variable. Therefore, the publication received the rating of C (inconclusive).

Average Publications (Rating 5-6)

Koenig, McCullough, and Larson (2001) rated six (27%) studies as average in their overall design. Of these studies assessing religious denomination and positive mental affect, four examined well-being, one assessed hope and optimism, and one studied self-esteem. The general conclusion reported from four of the studies is that religious denomination is not associated with positive affect. The remaining studies report that being affiliated with a religion serves as a beneficial aid to greater positive affect. The findings, whether positive or those declaring no association were reported in studies examining college students (Aycock and Noaker, 1985; Frankel and Hewitt, 1994), bereaved parents (Bohannon, 1991), religious nuns and female clergy (Magee, 1987; Rayburn, 1991), and elderly community dwelling adults (Tellis-Nayak, 1982).

The studies presenting positive findings between religious denomination and well-being conducted cross-sectional designs. For instance, Frankel and Hewitt (1994) examined 299 university students divided into two distinct groups. The first group consisted of members of a number of Christian clubs or faith groups, known as the affiliated group (N= 172). The second group, the non-affiliated group (N= 127), consisted of students enrolled in first and second year sociology courses. The research instrument employed was a self-administered questionnaire that contained questions on demographics, stress, mastery, self-esteem, psychological and physical health, use of health care resources, friendship patterns, beliefs, and values, and religious practice.

The questionnaire contained measures including psychological well-being, assessed by the Bradburn Affect-Balance Scale (Bradburn, 1969), mastery and self-esteem (Pearlin and Schooler, 1978), and a question concerning satisfaction with life. In the religious sample, the participants generally belonged to the Anglican denominations, Roman Catholic denomination, United Church denomination, Lutheran denomination, and Christian Reformed denomination.

Results reported using a t-test statistic show there to be differences between the two groups in relation to the Bradburn Affect-Balance Scale (Bradburn, 1969), both the positive and negative affect components of the scale. It can be observed that the affiliated group had higher scores on positive affect (M= 11.72) and lower scores on negative affect (M= 8.25) than the non-affiliated group (M= 10.96 and M= 9.34, respectively). These differences proved to be significantly significant for both positive and negative affect ($p < .001$). Additionally, it was reported that there were significant

differences between the groups in terms of satisfaction with life ($p < .001$). However, there were no significant differences with mastery or self-esteem. Although, the results from this study state that being in an affiliated group leads to greater positive effect, the analysis performed was zero-order. Consequently, the analysis did not control for potential confounding variables. For example, the results showed that there were many other differences between the two groups in relation to perceived health and physical health, that is, the affiliated group typically reported higher scores on perceived health ($p < .01$), less emergency room visits ($p < .05$), physician visits ($p < .05$), dentist visits ($p < .05$), and fewer days in hospital ($p < .001$). In addition, significant differences were found between the groups on number of stressful events ($p < .01$) and average level of stress per event ($p < .001$). If the analysis controlled for these variables then it could be possible that the significant differences between the groups, in relation to affiliation, would become weakened or non-significant.

Rayburn (1991) examined 254 religious females, consisting of 51 Roman Catholic nuns, 45 female Reform rabbis, 32 Episcopalian priests, 45 United Methodist clergy members, 45 Presbyterian clergy, and 36 seminarians. Subjects were aged between 24 and 75 years old. The subjects were matched on age, years of work experience, and pulpit duties. Results showed that Catholic nuns experienced less stress from work ($p < .001$), however, this analysis was uncontrolled.

The results presented from the above studies once again come from research that is cross-sectional and with the majority of the research performing zero-order analyses. Consequently, not incorporating controls in the analysis. The published studies in this section are not of adequate standard to provide enough evidence to establish if there is

enough support to suggest a relationship between the two variables, religious denomination and well-being. Therefore, the studies enter the inconclusive category and receive a rating of C.

Good Publications (Rating 7-8)

The vast majority of the studies (52%) in this category received the rating of seven to eight. These eleven empirical investigations are considered to be good studies in terms of their overall design. The general conclusion revealed by these studies is that there are no statistically significant association between religious denomination and positive affect. The studies reporting no association between the two variables employed high school and college samples (Bahr and Martin, 1983; Jensen, Jensen, and Wiederhold, 1993), adults residing within the community (Feigelman, Gorman, and Varacalli, 1992; Gee and Veevers, 1990; Mesienhelder, 1986; Russo and Dabul, 1997), and bereaved adults (Sherkat and Reed, 1992). However, few of the studies, which reported positive effects, included investigations between religious and non-religious kibbutz members (Anson, Carmel, Bonne, Levenson, and Maoz, 1990), between participants involved in “New Age” healing groups, Charismatic healing groups, and primary care medical patients (Glik, 1986), and for fundamentalists church members in comparison to church members involved in “moderate” or “liberal” denominations (Sethi and Seligman, 1993, 1994).

The general conclusion that can be postulated from the evidence provided by these studies is that religious denomination has a mixed relationship with well-being, depending on what sample group is investigated. In terms of the systematic review, the evidence provided by the studies would not be accepted by A-Levels-of-Evidence strategy as the publications cited by Koenig, McCullough, and Larson (2001)

employed a cross-sectional methodology, hence, presenting one of the flaws that prohibits inclusion under the regulations of the review process. Therefore, the publications receive a rating of C (inconclusive). However, within this subsection, the reviewers cited two studies that conducted a longitudinal methodology. One of the studies cited revealed the positive effects that religious denomination had on well-being (Glik, 1986) and one of the studies revealed that there was no association between the two variables (Russo and Dabul, 1997).

The longitudinal study showing the positive affect that denomination may exert on health; stated that individuals involved in Fundamentalist religious traditions generally report higher levels of positive affect. Conversely, whenever controls are added to the analysis the relationship between denomination and well-being generally drops. For instance, the theory that those from the Fundamentalist tradition experiences higher levels of psychological well-being, was explored in a study conducted by Glik (1986). The study performed by Glik (1986) was an exploratory comparison group study of participants in two spiritual healing practices and one from a medical care practice. The spiritual healing practices from which volunteers agreed to take part in the study came from 'New Age' or metaphysical (N= 93) and charismatic, largely Pentecostals (N= 83). The third sample, the primary care medical patients (N= 137), who were regular utilisers of the medical practice, were employed as the comparison group.

The subjects from the study were interviewed three times over 6-months, assessing socio-demographic and social factors, measures of physical, psychological, social, and behavioural health, religious attitudes and beliefs, and healing ritual behaviours.

The results from time 1 indicated that general well-being was significantly higher in the charismatic healing group, lower in the New Age healing group, and lowest in the primary care medical patients group ($p < .001$). These differences could not be explained due to the physical health of participants, as there were no significant differences in the severity of physical illness between the groups. Consequently, when the analysis adjusted for the covariates the significant differences between group and general well-being did not persist. The study commented on revealed the importance of extraneous variables, which may explain away the association between religious denomination and well-being,

The second longitudinal study cited within this subsection reported that there was no association between denomination and well-being. Russo and Dabul (1997) conducted a 7-year prospective cohort study on the relationships of abortion and childbearing to well-being on 1,189 black and 3,147 white women. It was reported that the 1987 well-being of women reporting a religious affiliation in 1979 ($M = 33.45$) did not differ from that of women reporting no affiliation ($M = 32.88$), $F(5, 4150) = .59$, ns. This result persisted, regardless of race.

The longitudinal studies presented within the subsection reveal mixed findings, with one study reporting a weak positive association between religious denomination and well-being, and the latter study presenting no association between the two variables. In examining the studies under A-Levels-of-Evidence strategy it can be proposed that the studies would be considered of adequate standard to be evaluated under the strategy. The studies performed longitudinal designs, controlled for potential confounders, adequately measured the religion and health variables, and performed

statistical analyses. The empirical study performed by Glik (1986) will be evaluated under the mediated model, as it did not control for protective factors (i.e., physical functioning, negative life events). The study by Glik (1986) did not contain any of the flaws under the evaluation category, thus therefore receives a rating of A (conclusive) + (positive result). The latter study performed by Russo and Dabul (1997) in a similar manner as the previous study will be evaluated under the mediated model and not the independent model, as it did not control for protective factors. The study by Russo and Dabul (1997) did contain one of the flaws under the evaluation category; it did not provide precise measurement of the well-being variables. The well-being variable was simply identified as self-esteem without any further description of what specifically was measured. Therefore, the study would be rated as B (generally sound), although the relationship was non-significant.

Excellent Publications (Rating 9-10)

Few studies received the rating of very good to excellent (14%). This equates to three studies out of twenty-one assessing the relationship between religious denomination and health. The findings reported from these studies stated that certain denominations were associated with higher levels of positive affect. This supposition has been supported in studies investigating the life satisfaction in community dwelling adults (Ellison, Gay, and Glass, 1989), life satisfaction in a community sample of Black adults (Ellison and Gay, 1990), and subjective well-being in an adult sample residing in the community (Ellison, 1991).

One of the studies conducted by Ellison, Gay, and Glass (1989) examined the relationship between religion and life satisfaction in a predominately white sample.

The researchers utilised data from a national random sample of persons aged 18 or over in the United States. This cross-sectional study consisted of 1,500 randomly selected individuals from the United States.

Life satisfaction was assessed by four-domain specific measures (single-item measures assessing finances, health, family, and friends). Religious denomination represented one aspect of the religious commitment measure. The religious denomination groups included: Catholic, Jewish, Southern Baptist, other Baptist, Lutheran, Presbyterian, Episcopal, Methodist, and other Protestant.

Results from regression analyses reported that the control variables (i.e., being married, less traumatic events, being female, occupational prestige, income, being white, and living in outside urban areas) were related to greater life satisfaction ($R^2=13\%$). The initial addition of the religious denominations, after sociability measures, into the analyses increased the variance by 1%, with all the denominations remaining significant, except Jewish and Presbyterian. However, with the addition of the remaining religious commitment measures, only Baptist religious affiliates remained related to life satisfaction ($p < .05$). This suggests the relative primacy of those religious factors that are not specific to any one particular denomination or sect, with the exception of the Baptist affiliates (Ellison, Gay, and Glass, 1989). At the final analysis, the full model only accounted for 18% of the variance in life satisfaction. In the final model, it would appear that the best predictors of life satisfaction were being married ($\beta = .14$, $p < .01$), income ($\beta = .13$, $p < .01$), and not being white ($\beta = -.12$, $p < .01$), these variables remained consistent throughout all six models. In addition, adding in devotional intensity emerged as a significant predictor ($\beta = .12$, $p < .01$).

In response to the findings that blacks from the previous study were at an elevated risk from lower levels of life satisfaction, Ellison and Gay (1990) conducted a similar study investigating African-Americans. This study produced comparable results to the former community study. The study utilised data from the National Survey of Black Americans conducted by the Survey Research Center during 1979-1980 (Jackson and Gurin, 1987). The sample consisted of 2,019 respondents aged 18 and older ($M= 43.15$, $SD= 17.71$).

The questionnaire included a measure of life satisfaction and denominational preference (i.e., Baptist, Methodist, Catholic, Fundamentalist, mainline Protestant, and other), which measured one aspect of religious commitment. In addition, Ellison and Gay (1990) included a variety of demographic variables (i.e., age, gender, family income, educational attainment, marital status, and rural/urban residences), personal stress, which included information on physical hindrances and interpersonal or institutional aggravators, and items measuring social networks. These items were employed as controls in the analyses between the religion variables and the well-being variable. An analysis of the findings revealed that the results from this study was comparable to the previous research conducted by Ellison, Gay, and Glass (1989). It was reported that the majority of the variance accounted for within the model ($R^2= 13.2\%$) was predicted from the demographic and social factors (i.e., age, residence, marital status, stress, and family closeness). The inclusion of religious denomination increased the explanatory power of the model modestly ($R^2= 13.9\%$), with all denominations demonstrating statistical significance with the exception of the 'other' group. The contributing power of religious participation increased the variance by

.3%, whilst adding in subjective religiosity and frequency of prayer did not increase the predicative power of the model. Once again, this data set demonstrates the unassuming presence of the religious variables.

According to the authors, the results from the study offer support for the social cohesion hypothesis. The social cohesion hypothesis postulates that religion benefits health by providing a sense of coherence and meaning, so the individuals understand their role in the universe, the purpose of life, and develop the courage to endure suffering (George, Larson, Koenig, and McCullough, 2000).

Ellison (1991) validated this supposition and reported from analyses that religion was an important source of existential coherence. In addition, Ellison (1991) claimed that the research demonstrated that persons with liberal, non-traditional, and nondenominational Protestant ties report significantly greater life satisfaction than unaffiliated individuals. These hypotheses were explored in an investigation examining the relationship between religious involvement and subjective well-being. The data for the study was taken from the General Social Survey (Davies and Smith, 1989), producing a national community representative sample (N= 997). This particular survey consisted of questions on religious socialization, belief, and practice. In addition, information on demographic and background characteristics, social integration, recent life crises, and subjective assessments of life quality were included.

From the data set, questions were utilised that assessed two dimensions of subjective well-being, affective and cognitive. Items that measured religious involvement included denominational preference, frequency of attendance at religious services,

two items on divine interaction, and three items measuring existential certainty. The background variables include aged, education, family income, race, gender, marital status, region, and rural-urban residence, these factors made up the demographic characteristics. In addition, stress was measured through the number of traumatic life events (i.e., divorce, unemployment, and hospitalisation/disability) that happened during the year preceding the survey. Also considered was the frequency of interpersonal interaction and total number of memberships in secular voluntary associations, these items measured evidence of social interaction and support.

Results from the analyses on the cognitive component of subjective well-being (life satisfaction), showed that demographic characteristics explained 10.9% of the variance for life satisfaction. However, when the religious variables were added to the model the variance was increased to 16%, with liberal Protestants ($p < .05$), non-denominational Protestants ($p < .001$), divine interaction ($p < .05$), and existential certainty ($p < .001$) appearing as significant predictors, with existential certainty emerging as the most significant predictor ($\beta = .131$). The final model confirmed the impact that social factors, such as increasing age, more years in education, more income, being married, less traumatic events, and more voluntary association memberships has on life satisfaction, with the best predictor being education ($\beta = .849$, $p < .001$).

The additional variance accounted for in life satisfaction above the demographic items by the inclusion of the religious predictors in the model was far from enhanced greatly ($R^2_{\text{change}} = .051$) with existential certainty and non-denominational Protestant emerging

as the best predictors. It can be speculated from the studies mentioned that the best predictors of life satisfaction include demographic variables. This proposition was validated by subsequent research (e.g. Reed, 1991).

The analysis that concerned the second component of subjective well-being, personal happiness, reported on similar results to the previous analysis. Findings revealed that the demographic variables were the best predictors of personal happiness, with income, marital status, fewer traumatic events, and voluntary association memberships being statistically significant. These variables in model 1 accounted for 10.3% of the variance in personal happiness. The addition of the religious variables did not enhance the variance greatly ($R^2_{adj} = .127$). In model 2, it was observed that with the addition of the religious items, it was observed that being Catholic was negatively related to personal happiness. Existential certainty, as before, was positively associated with personal happiness. The best model for this data set explained 13.1% of the variance, and included income, being married, fewer traumatic life events, involvements with voluntary associations, being Catholic (negative), and existential certainty showing statistically significant relationships with personal happiness. The single best predictor that emerged was fewer traumatic life events ($\beta = -.519, p < .001$).

From the analysis it can be speculated that existential certainty, a measure of how often doubts about the respondent's religious faith have been caused by evil in the world, conflicts between faith and science, and the feeling that life really has no meaning, is a superior measure of religiosity than the more popular measures of denomination, church attendance, and frequency of prayer. However, it could be

argued that the items measuring existential certainty are tapping into a construct measuring optimism or hope.

The studies presented associations between religious denomination and positive affect, albeit a weak association, they were not without their limitations. For instance, the operationalisation of the life satisfaction concept in Ellison, Glass, and Gay (1989) and Ellison and Gay (1990) was inadequate. In all the studies presented within this category, the measurement of life satisfaction was assessed through four-domain specific measures (single-item measures assessing finances, health, family, and friends) the responses to these items were summed together to provide a composite score of general life satisfaction. However, it can be speculated that the items could be measuring quality of life rather than life satisfaction. Life satisfaction is complex concept to define, which presents difficulties in its operationalisation. Therefore, researchers are far from forming a consensus regarding the conceptualisation and measurement of this variable and the strength of an association between the two depend largely on how the variable is measured (Chamberlain and Zika, 1988).

It can be assumed by observing the findings from the studies that the relationship between religious denomination and well-being is one that is positively associated, with certain denominations having positive scores on measures of well-being. However, the typical foundations for this conclusion come from cross-sectional designs and under the systematic review process the evidence from the publications cannot be employed as the methodology is of inadequate design. Therefore, the three publications under A-Levels-of-Evidence approach are rated as C (inconclusive).

After re-evaluating the studies within this category, it can be observed that some have shown that denomination is associated with positive affect, though weakly. However, the majority of the published studies have demonstrated that there was no association which existed between the two variables. For instance, Aycock and Noaker (1985) concluded that evangelical Christians did not report superior levels of self-esteem in comparison to the general population. Aycock and Noaker (1985) examined the self-esteem levels of 351 evangelical Christians from college and church settings and 1115 general volunteers comprised of students, administrators, and government employees. Self-esteem was measured by the Coping Resources Inventory for Stress (Matheny, Curlette, Aycock, Pugh, and Taylor, 1981). Results from the analysis of variance showed that comparison of the self-esteem scores differed significantly, ($F(11,1494)=12.16, p < .001$); however, these differences were attributable to the subgroups amongst the general population. It was found that educational attainment produced the significant differences, not religious affiliation.

Koenig, McCullough, and Larson (2001) state that 'In the vast majority of the studies, religious involvement was positively associated with greater well-being' (p.117). Indeed, these authors do not comment specifically on religious denomination, although, observing the catalogue it would appear at first glance that religious denomination has positive effects on positive affect. Conversely, in closer examination of the studies the overall conclusion, which can be commented on, is that religious denomination is generally unrelated to positive mental health, such as well-being, self-esteem hope and optimism. It would appear that in the studies reporting positive associations, the most significant predictors of positive affect are

demographic characteristics, such as marital status, voluntary association memberships, traumatic life events, and income (Ellison, 1991).

In examining the publications assessing religious denomination and subjective well-being, which were cited by Koenig, McCullough, and Larson (2001), it would seem that the relationship is mixed, with some studies presenting positive associations and some presenting no associations. However, utilising an objective strategy to evaluate the studies it is possible to establish which of the studies meet minimally acceptable methodological standards that could possibly depict the direction of the relationship.

From the twenty-one publications cited by Koenig, McCullough, and Larson (2001), only two of them meet the minimal methodological standards to be evaluated under the systematic review approach. In terms of evaluating the strength of the hypothesis; denomination increases subjective well-being. In examining the ratings provided by the evaluation from the systematic review process, it appears that there are enough high quality studies to infer a conclusion based on the data from the studies. Two studies from this that section assessed the relationship between religious denomination and depression are considered to meet the minimal methodological standards to provide an evaluation of the strength of the hypothesis that religious denomination is one factor attributable to an increase in subjective well-being. The two publications were either rated in the A (positive) category and the B (ns) category. Therefore, due to the study that demonstrates conclusive evidence for the hypothesis that religious denomination is attributable to an increase in subjective well-being, the data reaches the level of *some evidence*.

Conclusion

The belief projected by the majority of researchers assessing this domain, is that certain religious groups are protected either from or at risk from suffering from positive or negative mental health disturbances. Specifically, an element, which is consistently communicated within the literature is those from Jewish traditions appear to be at an elevated risk from suffering from depressive symptomology. The opinion that those from Jewish populations are at a higher risk from suffering of depressive symptomology has been reported in studies dating back to the 1880s. Sauna (1992) performed a review on a large group of cross-sectional epidemiological and clinical studies related to the prevalence of mental disorders, including depression that included a Jewish sample (e.g., Fernando, 1978; Figelman, 1968; Cooklin, Ravindran, and Carney, 1983; Flics and Herron, 1991; Malzberg, 1973). Sauna (1992) concluded from the literature that people from Jewish descent appear to be associated with higher depressive disorders. Sauna (1992) also commented that Jews tend to have higher levels of depressive disorder than non-Jews, and that this finding remained prevalent when psychiatric symptoms of admissions to psychiatric hospitals were examined. This finding that Jews from a psychiatric population were at an elevated risk from depressive symptomology has been reported in the United Kingdom (Fernando, 1978; Cooklin, Ravindran, and Carney, 1983) as well as in the United States (e.g., Flics and Herron, 1991; Malzberg, 1973).

In response to this claim made by Sauna (1992), Levay, Kohn, Golding, and Weissman (1997) point out that most of the studies under review use clinical samples rather than generating data from the general population. Consequently, using data from such a restrictive population may lead to a variety of biases that may jeopardise

the validity of the findings. Furthermore, it has been documented that those from Jewish traditions are more inclined to seek psychoanalytic psychotherapy in comparison to those from other denominations (e.g., Weintraub and Aronson, 1968), leading to an overrepresentation of the Jewish population in clinical samples. As a result, it could be speculated that the interpretation of results from clinical populations may be skewed and unreliable.

However, it emerged through subsequent empirical studies, which utilised community samples in opposition to the more limited clinical samples that this general conclusion is validated (e.g., Kennedy, Kelman, Thomas, and Chen, 1996; Ross, 1990; Yeung and Greenwald, 1992). These community-based epidemiological studies, report that on average Jews appear to have approximately a two-fold increased risk for suffering from major depression compared to members of other religious groups. However, as discussed previously these studies are fraught with methodological and conceptual limitations that prevent justifying this general conclusion.

The review by Koenig, McCullough, and Larson (2001) presents a picture in which Jews would appear to be at an elevated risk from symptoms associated with depression. However, in closer examination of the studies assessing Jewish populations, it would can be viewed that they are fraught with methodological flaws that require addressing before such a conclusion can be validated. For example, the majority of the research does not meet the minimal methodological standards to substantiate this conclusion.

When looking at the general conclusion that religious denomination offers a protective factor against depression, it can be seen that there is enough evidence available to validate this conclusion. Four studies from the section examining religious denomination and depression offered conclusive findings in regard to this hypothesis. Although, one of the studies provided support that was negative (Kennedy, Kelman, Thomas, and Chen, 1992), the remaining three studies provided positive support for the hypothesis (Idler and Kasl, 1992; Koenig, Cohen, Blazer, Pieper, Meador, Shelp, Goli, and DiPasquale, 1992; Miller, Warner, Wickramaratne, and Weissman, 1997). The findings from the three studies positively supporting the hypothesis reached the level of *persuasive evidence*.

In relation to the cited publications which examined the relationship between religious denomination and anxiety, it can be initially suggested that the association is one that is generally mixed. However, when the publications are subjected to the systematic review process, it can be observed that none of the publications meet the minimal methodological standards to be evaluated for the strength of any hypothesis. The publications cited by Koenig, McCullough, and Larson (2001) being rated in the C category under the regulations of the systematic review process, this signifies that the results from the studies reach a level of *insufficient evidence*.

Examining the publications that investigated the relationship between religious denomination and subjective well-being, it can be viewed that an extensive amount of research was conducted within the domain. With the majority of research suggesting that religious denomination is associated with increased levels of subjective well-being. However, when the publications from the section was scrutinised under the

systematic review process, it was established that the majority of the empirical research did not meet the methodological standards to be evaluated. However, within one of the subsections, two studies (Glik, 1996; Russo and Dabul, 1997) did meet the standards and it was demonstrated that there was *some evidence* to support the hypothesis.

Within the three mental health issues that have been substantially investigated by empirical researchers, it was observed that the majority of the existing studies do not meet the minimal methodological standards to be assessed under an objective evaluative process. This process states whether a relationship between religious denomination and anxiety, depression, and subjective well-being, is likely to exist. However, few studies within the sections examining depression and subjective well-being did provide adequate, methodologically sound studies to evaluate the strength of any hypothesis considered. In considering the conclusions surrounding each mental health issue, stated by Koenig, McCullough, and Larson (2001) it would appear that the review is generally overly optimistic in observing the relationship between religious denomination and mental health. Similarly, the work of Sloan and colleagues (e.g., Sloan and Bagiella, 2002) tends to be overly pessimistic. Rather it would seem that there is a dearth of high-quality research conducted on the subject; although, the evidence to date warrants continued and careful investigation.

2 Religious Denomination and Physical Health

Introduction

The main premise within this current section is to examine the empirical research concerning the effect religious denomination has on physical health. The review conducted by Koenig, McCullough, and Larson (2001) extensively assessed this topic with a range of physical health issues. Hence, the intention of this chapter is to present their findings and re-examine the studies that these reviewers considered to be of adequate value within this area. Consequently, this chapter will evaluate critically the research conducted on physical health (heart disease, hypertension, cancer, and all-cause mortality). It will be shown that a relationship exists between these two variables; however, it is probable that the relationship is not as warranted as that discussed in the review (Koenig, McCullough, and Larson, 2001). Therefore, by incorporating the objective criteria enlisted by the systematic review, it is possible to perform an evaluation of the evidence within this domain.

Cancer

Cancer is an illness that provokes fear in many people. It is an illness that can affect any organ within the human system, potentially leading to mortality. It has been estimated that the lifetime likelihood that a man in the United States will contract cancer is 50%, and for woman it is approximately 33%, with the figures being similar for individuals from the UK (American Cancer Society, 1999; CancerHelp UK, 1996).

The relationship between religion and cancer risk and mortality has been the focus of numerous published articles by epidemiologists and public health researchers over several decades (Levin and Schiller, 1987). Generally, it has been found in earlier studies that those individuals particularly from religious groups that are characterised by doctrinal orthodoxy and behavioural conformity are at a reduced risk of cancer (e.g., Enstrom, 1975; Philips, 1975). As a result, suggestions have been made that members of different religious groups have varying risks of developing cancer.

Levin and Schiller (1987) reviewed research looking at religion and health. Within their review, they also examined the association in relation to cancer and religious denomination. They reported that the outcomes from studies showed that results varied across specific sites. However, significant protection overall was found among Hutterites (e.g., Martin, Dunn, Simpson, Olsen, Kemel, Grace, Ellas, Sarto, Smalley, and Steinberg, 1980), the Amish (e.g., Hammam, Barancik, and Lilienfeld, 1981), Mormons (e.g., Enstrom, 1978), and Seventh Day Adventists (e.g., Lemon, Walden, and Woods, 1964; Philips, Kuzma, Beeson, and Lotz, 1980; Berkel and de Waard, 1983). In terms of those from the Jewish tradition, researchers would report that cancer mortality rates fall somewhere below that of Catholics and above that of Protestants (e.g., Bolduan and Weiner, 1933; Seidman, 1966). Other evidence provides a contrasting story to this postulation (e.g., Horowitz and Enterline, 1970; MacMahon, 1960).

The general conclusion is that there is an association between individuals from certain religious denominations and the risk from suffering from certain types of cancers. This supposition can be validated with more recent studies examining the association.

However, these studies are fraught with much of the same conceptual and methodological limitations that are experienced by empirical research examining the effects of religious denomination on mental health.

As with the empirical studies examining the effects that denomination had on facets of mental health, Koenig, McCullough and Larson (2001) listed and rated empirical studies that examined the effect that religious variables exerted on cancer risk and cancer mortality. The reviewers listed fifty-three studies that utilised some form of the religiousness variable, either as a behavioural, affiliated, beliefs, or attitudinal construct. From these fifty-three studies, the majority (42) of the studies examined denomination, in particular, comparing different denominations in relation to cancer risk and morality.

Weak Publications (Rating 1-4)

In relation to the weak category, Koenig, McCullough, and Larson (2001) listed few studies in this subsection. From the forty-four studies catalogued in the review, only nine (21%) received ratings from one to four, with the vast majority of the studies appearing at the upper end of the rating scale. The results from these studies revealed mixed findings. The majority of the studies in this category assessed the cancer risk and morality of Jewish individuals in comparison to Christian traditions. For instance, it would appear that the findings from these set of studies within this subsection revealed that Jewish males have increased risk of contracting and dying from cancer of the liver, 'urinary organs', rectum and pancreas. Jewish females are at greater risk from cancer of the breast, lung, large intestine, liver, pancreas, lymphatic tissues, haematopoietic tissue (Hoffman, 1932; King, Diamond, and Bailar, 1965;

Newill, 1961). However, Jewish men have less risk from suffering cancer of the respiratory system and digestive system and cancer of the 'reproductive organs' is less common in Jewish women (Hoffman, 1932; Horowitz and Enterline, 1970; King, Diamond, and Bailar, 1965; Newill, 1961; Rosenwaike, 1984). Studies conducted in Australia have reported that there are no differences in large-bowel cancer risk (Kune, Kune, Watson, and Bahnson, 1991) or survival from colorectal cancer mortality (Kune, Kune, and Watson, 1992) in a Jewish sample compared to members from other affiliations. In addition, an early report conducted at the beginning of the 19th century in London (Correspondent, 1902) reported that non-Jews had a greater chance of suffering from cancer than Jews; however, no statistical comparison was made and socio-economic class was not controlled for. In a study conducted in Bombay, Jussawalla and Jain (1977) reported that Zoroastrian women had greater reports of breast cancer than women from other affiliations did. It would seem that the majority of the studies concluded that members of the Jewish religion have greater likelihood of suffering from certain forms of cancer risk and cancer mortality.

On initial examination of the studies assessing religious denominations and cancer risk and mortality, it would seem that certain groups, in particular Jews, are at an elevated risk from suffering from certain cancer sites. However, the studies are mostly cross-sectional; hence, it is not possible to infer any concrete conclusions. The publications, which present such conclusions come from research that is methodologically weak, in that they incorporate cross-sectional designs and generally provide such conclusions without the employment of controls or with the employment of low-order controls, such as age and sex. Utilising high-level controls, such as health behaviour and socio-economic status, may explain the observed relationship.

Regarding these flaws, the publications in the subsection would be rated in the category C, which represents evidence that is flawed and inconclusive. The prospective cohort studies conducted by Kune, Kune, Watson, and Bahnson (1991) and Kune, Kune, and Watson (1992) controlled for multiple confounders in their analysis.

The first study performed by Kune, Kune, Watson, and Bahnson (1991) primarily examined personality as a risk factor in large bowel cancer. The study consisted of 637 cases and 714 controls derived from a large, comprehensive, population-based clinicopathological and epidemiological study of colorectal cancer incidence, aetiology, and survival. The cases were all histologically confirmed new cases of colorectal adenocarcinoma. In order to minimise possible selection biases or exclusion bias, an extensive investigation was made comparing age, sex, country of birth and religion characteristics of cases and controls that were included in the study and these were compared to those not included in the study. The sample was administered two questionnaires. The first included data on age, sex, country of birth, religion, current and past illnesses, operations, medications, bowel habit, parity and hormonal factors, family history data and psychosocial factors (object loss and unhappiness, cancer prone personality, reactions of anger). The second questionnaire included information on diet, alcohol intake, and tobacco consumption. Results from the analysis did not assess the relationship of religion to cancer risk or mortality specifically. However, religion was assessed as a confounder variable, with reports that there were significant differences between cases and controls in cancer personality scores, which were independent of religion.

Although, this latter study met almost all the essential criteria to be evaluated under the systematic review process, the authors, Kune, Kune and Watson (1992), utilised the same cohort data for a study that they performed in the following year. Subsequently, the study conducted by Kune, Kune, Watson, and Bahnson (1991) would be omitted from the systematic review as the data from the same cohort was utilised twice and under the regulations of A-Levels-of-Evidence strategy the study, which conducted the longest study period is employed.

The study conducted by Kune, Kune, and Watson (1992) examined the risk of suffering from colorectal cancer and cancer of the large intestine. The aim of the study was to investigate the effect of family history of cancer, religion, parity, and migrant status on survival in cancer of the colon. The study consisted of 705 histologically confirmed new cases of colorectal adenocarcinoma during a 12-month period. Each of the respondents provided information on age, sex, cancer site, type of cancer, degree of cell differentiation, and the presence of perforation, as factors of interest with respect to survival. The factors examined in relation to survival included migrant status, religion, family history of colorectal cancer, number of children and age of birth of first child.

Findings from the analysis showed that the respondents ranged in age from 19 to 93 years ($M= 66$, $SD= 11$). In relation to the religion factor, results from the analysis showed that religion was not associated with cancer survival. The factors that were related to cancer survival over the 5-year period were age (CI 1.00-1.03, $p < .05$), cancer site (CI 0.66-1.16, $p < .01$), in which survival was worse for rectal cancer, than

colon, cancer, and better cell differentiation (CI 0.56-0.93, $p < .01$). Religion had no influence on survival in any of the stages of colorectal cancer.

The study performed by Kune, Kune, and Watson (1992) meets the criteria for inclusion to the review process; that is, the study controlled for confounders, is of a longitudinal design, measured religion adequately, performed statistical analyses, and utilised a preliminary cohort. The rating that this study would receive under the review process is a rating of B (ns), both for the mediated and independent models, indicating that the evidence provided from the research is generally good. It does not receive the rating of A (conclusive) as it has one flaw; it does not precisely measure religion, the variable was merely identified as 'religion', with no further description of what was being measured.

Average Publications (Rating 5-6)

Half of empirical research (50%) was cited in the average category (rating 5 to 6). The majority of these 21 studies (50%) reported that certain denominations are at an elevated rate of suffering from cancer risk or cancer mortality. The groups that appeared to be at greater risk of certain cancers and mortality were Jewish, non-Seventh Day Adventists, and non-Mormons. Findings showed that Jewish male individuals were more at risk from cancers of the rectum, whilst females were at a greater risk of breast and lung cancer mortality (Versluys, 1949; Wolff, 1939). Additionally research reported supportive findings that Jewish males have less risk and mortality from cancers of the lung (Greenwald, Korn, Nasca, and Wolfgang, 1975; Herman and Enterline, 1970; Seidman, 1966, 1970, 1971). These findings would appear to validate the general findings from the previous section. However,

few studies reported opposite findings to those in the previous section. For instance, King, Diamond, and Bailar (1995) reported that Jews were less likely to suffer from Leukaemia and Hodgkins Disease; this is in direct opposition to that reported by MacMahon and Kroller (1957) and MacMahon (1960), in which it would appear that Jewish men and women have greater incidences of Leukaemia risk and mortality. However, the period and region in which the sample was examined could have confounded this difference in results.

Former studies reporting positive associations between Judaism and cancer failed to control for important covariates, such as lifestyle behaviours (i.e., smoking, alcohol use, family history, and dietary behaviour). Studies within this subsection that incorporated multiple variables in the analysis, such as family history revealed that ovarian and breast cancer risk and mortality was not associated with members of the female Jewish tradition (Steinberg, Pernarelli, Marcus, Khoury, Schildraut, and Matchbanks, 1998; Tonio and Kato, 1996). These findings could represent the importance of genetic factors in cancer.

Correspondingly, studies that showed individuals outside Mormon or Seventh-day Adventist religious traditions, have a greater likelihood of suffering from cancer risk and mortality of the respiratory system (Fraser, Beeson, and Philips, 1991; Lyon, Gardner, Klauber, and Smart, 1977; Philips, 1975). These studies did not incorporate controls into their; however, when researchers utilised low-order controls, such as age and sex, they reported similar findings (Jarvis, 1977; Lemon, Walden, and Woods, 1964; Lemon and Kuzma, 1969; Lyon, Klauber, Gardner, and Smart, 1976). Conversely, when smoking and alcohol behaviour were included in the analyses the

differences reported between Seventh-day Adventist's and the controls (non-Seventh-day Adventists), dropped substantially, usually to nonsignificance (Jensen, 1983).

The majority of the studies within this subsection examining the relationship between denomination and cancer risk and cancer mortality do not meet the criteria to be included in the systematic review process. The evidence that was presented was deemed to be inconclusive, and thus the studies receive a rating of C. However, few studies within this subsection do meet the criteria for inclusion. The typical characteristics of these publications is that they present research with a longitudinal design, incorporated multiple controls, provided adequate measurement of the religion variable, performed statistical analysis, and employed a preliminary cohort.

The first study performed by Jensen (1983) met the minimal methodological standards to be evaluated under the review. Jensen (1983) investigated cancer risk among Danish male Seventh-day Adventists in comparison with other temperance society members. The study consisted of 781 Adventists and 808 non-Adventists who did not drink alcohol. Cancer morbidity and mortality status were established using municipal records, the National Central Person Registry, and National Central Death Registry. Results from the analysis revealed that Adventists had a standardised mortality ratio of 0.69 or 31% ($p < .05$) lower rates for all cancers, whereas the temperance members had standardised mortality ratios of 1.05. This result is not significantly different from one, so, is not statistically significant. However, when alcohol and tobacco related cancers were excluded (i.e., buccal cavity and pharynx, oesophagus, larynx, lung, and bladder), the standardised mortality ratio for Seventh-day Adventists rose from 0.69 to 0.93, a figure that was not significantly different

from one. Jensen (1983) revealed the importance of lifestyle behaviours, the analysis revealed how lifestyle factors influences chronic illnesses.

The second study performed by Toniolo and Kato (1996) assessed the relationship between Jewish religion and breast cancer in a prospective cohort study of 10,273 women enrolled between 1985 and 1991 for mammographic screening in New York City. Among Jewish women age 50 or younger with a family history of breast cancer, the relative risk was again higher in Jews than in members of other religious groups (2.33, 95% CI 1.35-4.02, adjusted for multiple risk factors). The authors found no increased risk among Jewish women without a family history or among those over age 50. The authors concluded that certain groups of Jewish women have a higher than expected rate of mutation in breast cancer gene.

The collection of published studies presented within this sub-section, are estimated to be of inadequate standard to be included under A-Levels-of-Evidence approach. Therefore, would receive a rating of C (inconclusive). However, one of the studies (Jensen, 1983) does meet the criteria for inclusion, and hence receives a rating of A (positive) for the mediated model and the rating of A (ns) for the independent model, meaning that the methodology of the publication is conclusive and supports the hypothesis positively. However, in relation to the independent model, support for the hypothesis is nonsignificant and it's the inclusion of a variety of protective factors, such as not smoking or drinking that explains the observed variance. The observed difference between the two models can be interpreted that the religion and cancer risk or mortality link was accounted for by risk factors, in this instance genetic influences.

The second study performed by Toniolo and Kato (1996) is of adequate standard to receive a rating from A-Level-of Evidence approach. Ratings for the mediated model, in which the impact of religion on health is evaluated regardless of whether or not such a relationship is mediated by established risk factors. That is, the strength of the evidence for a relationship is evaluated without control for any potential mediating factors such as healthy lifestyle, social support, and depression. Utilising the criteria for the first model, the published study does not receive a rating as the authors did not employ any controls for potential confounders such as socio-demographics, although age was controlled for this was not an adequate amount of potential confounders to exclude the effects of other socio-demographic characteristics. However, in terms of the independent model, the study would receive a rating of B (ns), signifying that the evidence provided by the study was generally sound but nonsignificant, due to the effects from family history of breast cancer.

Good Publications (Rating 7-8)

The group of studies rated as good (7 to 8), by Koenig, McCullough, and Larson (2001), are the second biggest in this category. Eleven (26%) studies examined the relationship between religious denomination and cancer, both risk and mortality. The general findings reported by the studies are consistent with the conclusions reported from studies in the former categories; Jewish women are at greater risk from breast cancer than non-Jewish women (Egan, Newcomb, Longnecker, TrenthamDietz, Trichopoulos, Stampfer, and Willett, 1996), Mormons and Seventh-day Adventists have less cancer risk and mortality than individuals outside these religious traditions (Enstrom, 1975, 1980; Gardner and Lyon, 1977; Lemon and Walden, 1966; Philips, Garfinkel, Kumza, Beeson, Lotz, and Brin, 1980; Philips, Kuzma, Beeson, and Lotz,

1980). Most of the studies controlled for age and sex, so, it was expected that the differential effect would disappear once multiple controls were employed. This finding can be observed in a study examining breast cancer among Seventh-day Adventists and non-Seventh-day Adventists (Zollinger, Philips, and Kuzman, 1984). Nevertheless, in two high quality studies this differential effect remained when multiple controls were employed (Enstrom, 1978; Dwyer, Clarke, and Miller, 1990). These findings may potentially imply that factors outside of doctrinally encouraged health behaviours are important elements in extending longevity. These studies were of a cross-sectional nature, which may indicate this postulation; however, it does not confirm it. In addition, it has been suggested that cancer may be attributable to ecological and environmental factors, such as chemicals used in agriculture and airborne pollutants in industrial settings have been recognised as cancer-causing agents (Stokes and Brace, 1988). Consequently, it is important to control for these variables in modelling the relationship between religion and cancer mortality rates.

Dwyer, Clarke, and Miller (1990) conducted a cross-sectional random study across 3063 counties in the United States examining the effects of religious concentration and affiliation on cancer mortality rates. The study was conducted in response to the scarce literature, which does not examine differences in mortality, by major denominations (Jarvis and Northcott, 1987), or for studies, that do not examine the possible effects of ecological and environmental confounders. Therefore, an approach that incorporates factors known to affect cancer mortality and that uses aggregate data will provide insight into the relationship between religion and health (Levin and Schiller, 1987).

The county-level data employed in the study was taken from a variety of sources. The data utilised in the study included specific forms of cancer, all malignancies combined, for cancer of the digestive system and respiratory cancer. Data on age, sex, and race, church attendance, denominational affiliation were also utilised. Controls for demographic, environmental, and regional factors known to affect cancer mortality were included explicitly in the model (Dwyer, Clarke, and Miller, 1990). The denominational categories included in the analysis were conservative Protestant, moderate Protestant, liberal Protestant, Catholic, Jewish, and Mormon. Non-communicants were employed as a reference group. The data utilised was taken from sources available from 1968 to 1970, 1971 to 1975, and 1976 to 1980.

Initial findings from the study reported that observing the socio-demographic structure of the community affects cause-specific cancer mortality rates. For instance, across the three time bands, it was revealed that conservative Protestants and Mormons had the lowest cancer mortality rates, whilst counties with the higher concentration of Jews or liberal Protestants had the highest cancer mortality rates. These findings remained consistent with the control of fifteen factors known to affect cancer mortality. However, it should be noted that using samples that practise strict doctrines that proscribe behaviours associated with risk might lead to misleading results, as individuals from these populations will report greater health.

Interpreting the evidence from the studies cited within this subsection, would read as equivocal; some studies reported on a relationship, whilst others reported no association. In terms of the objective review process, the evidence provided from the majority of these studies come from cross-sectional designs and cannot be entered

under the systematic review. Therefore, the empirical studies received a rating of C, which considers the evidence as inconclusive. However, there were a few studies within this subsection, which do meet the criteria for inclusion under the review strategy. For instance, the first study conducted by Philips, Kumza, Beeson, and Lotz (1980) was a 16-year prospective cohort study of 22,940 Seventh-day Adventists and a 12-year prospective cohort study of 112,726 non-Seventh-day Adventists. The subjects were both initially assessed in 1960, and where all over the age of 35 and white. For each subject person-years were calculated from January 1, 1960, until death, loss to follow-up, or termination of follow-up (July 30, 1972, for non Seventh-day Adventists; December 31, 1976, for Seventh-day Adventists). For the both groups, counts of deaths and persons-years accumulated over the follow-up period were stratified for age, sex, religion, smoking history, and thirds of an index reflecting degree of adherence to sixteen different health habits.

Results from the analysis revealed that greater numbers of Seventh-day Adventists did not smoke, did not consume meat, poultry, or coffee, than individuals who were non Seventh-day Adventists. In comparing the risk of dying from selected major cancers among the two groups, in terms of age-adjusted mortality ratios, it was revealed that there were significant differences in mortality ratios between Seventh-day Adventists and non Seventh-day Adventists for all cancer mortality, in male (0.60, $p < .001$) and females (0.76, $p < .001$). Similar results were found for lung cancer, in males (0.18, $p < .001$) and females (0.31, $p < .001$), and for colon-rectal cancer for males (0.62, $p < .001$) and females (0.58, $p < .001$). Therefore, it appears as if cancer deaths are reduced among Seventh-day Adventists. However, due to the health habits of the Seventh-day Adventist it could be possible that not smoking, consuming alcohol, and

leading mostly a vegetarian diet may decrease the risk of cancer; however, these factors were not controlled for in the analysis.

Zollinger, Phillips, and Kumza (1984) reported contrary findings to the previous study in a sample of 2,304 varying religious white women, including those reporting a Seventh-day Adventist affiliation. The study was a 30-year longitudinal study, which assessed the survival rate of individuals with cancer. The sample consisted of 282 Seventh-day Adventists, 1,675 non-Seventh-day Adventists, and 347 unknown religious patients with breast cancer in California. Results from the analysis revealed that for Seventh-day Adventists, breast cancer cases displayed a consistently better survival pattern than non-Seventh-day Adventists cases at various times after diagnosis. The cumulative probability of not dying of breast cancer after 10 years was 60.8% for Seventh-day Adventists and 48.3% for non-Seventh-day Adventists ($p < .05$). However, it was shown from the data that Seventh-day Adventists were diagnosed at an earlier stage, and matching variables (histologic type, malignancy grade, treatment type, age, and year at diagnosis) accounted for the small portion of the survival difference not attributable to earlier stage of diagnosis.

The evidence from these publications from this subsection, would present conclusions that are mixed. The first study (Philips, Kuzma, Beeson, and Lotz, 1980) stated that there was a significant relationship between the variables of religious denomination and cancer. However, the researchers did not employ high-level controls within their analysis; the analysis only controlled for age and sex. The second study stated that there was a significant relationship; however, the significant relationship was

diminished whenever potential confounders were utilised. Consequently, in terms of A-Levels-of-Evidence approach both of the studies would be included in the systematic review process, but only under the mediated model; not under the independent model as both the studies do not include the potential effects of protective factors (i.e., healthy lifestyle). Therefore, both studies would receive a rating under the regulations of the review process, of A (conclusive). The study by Philips, Kuzma, Beeson, and Lotz (1980), would be awarded a rating of A positive, and the study by Zollinger, Philips, and Kuzma (1984) would be awarded a rating of A non-significant.

Excellent Publications (Rating 9-10)

The last subsection within this category, only one study was perceived to be of an excellent standard by the reviewers (Koenig, McCullough, and Larson, 2001). The study performed by Enstrom (1989) examined the health practices and cancer mortality among active California Mormons. The prospective cohort study consisted of 9,844 religiously active Mormons, 5,231 high priests and 4,613 wives in California and compared them with a control population of 3,199 adults in Alameda County. The questionnaire consisted of items on basic demographic characteristics, major components of life-style, diet, medical history, and other health-related characteristics. The data incorporated parts of a 1974 Health and Ways of Living questionnaire used in the Alameda County by the Human Population Laboratory of the California State Department of Health Services (Enstrom, Kanim, and Breslow, 1986). Mortality follow-up data was obtained from 1980-1987 on all high priests in California.

Findings from the study revealed that Mormons had lower occurrences of cancer mortality than controls. For the high priests in the sample, standard mortality ratios were 47 compared to 100 for controls for all causes. When, three health practices (never smoked, got regular exercise, and got proper sleep) were included in the analysis the standard mortality ratio was increased to 48; although when two of the three health habits were included the mortality ratio was reduced to 46. For the females in the sample, the standard mortality ratio for all cancers changed from 72 to 56, with the adjustment of health practices. These findings reveal that the standard mortality ratio's generally declined with the adherence to more health practices. Findings appeared to reveal that the Mormon sample benefits from the protective effects against cancer due to their religious adherence.

A criticism of this study is that the researchers performed a case-control study, in which the subjects were matched in relation to age and sex. The strategy of matching ensures equivalence of groups only on the matched variables. As a result, the groups could differ with respect to social factors, such as socioeconomic status. However, the researcher controlled for a variety of health behaviours.

The evidence provided by the publication (Enstrom, 1989) revealed that there is a positive relationship between the two variables, religious denomination and cancer mortality. In terms of the systematic review process, the publication presented within this subsection could be rated and evaluated by A-Levels-of-Evidence approach, as it does not include the methodological flaws, which prohibit inclusion to the review. In examining the methodology under the mediated model, although the research did not control for all socio-demographic characteristics, the sample was matched on sex and

age; therefore, the study would be rated in the category A (positive). In relation to the independent model, the study utilises a few health practices and controls for them in the analysis, therefore, the study is rated in the category of A (positive). The non-discrepancy between the mediated and independent models denotes that health practices did not fully explain the observed results.

Overall, considering the publications cited by Koenig, McCullough, and Larson (2001) that examined the relationship between religious denomination and cancer risk and cancer mortality, it appears that few of them met the minimal methodological standards to be included within the systematic review. From the forty-four studies cited, only six of the published articles were considered to meet the minimal methodological standards. In evaluating the strength of the hypothesis, which states that religious denomination is associated with lower cancer risk and mortality, the six studies present evidence that is of adequate standard to make a judgement on the hypothesis. The first study cited in the weak category, which was performed by Kune, Kune, and Watson (1992), was rated both in the mediated and independent model, in the B category. However, the study offered no support for the hypothesis. In the second category, rating five to six, two studies met the minimum standards. The first study performed by Jensen (1983), was rated A (positive) in the mediated model, and as A (ns) in the independent model; indicating that it was the inclusion of health protective factors, which accounted for the observed association between the two variables. The second study within this category was rated in the category of B (ns). This rating indicates that the methodology for the study was generally sound; although, there was no support for the hypothesis. In the category, which was rated as good, there were two studies that were considered adequate standard to be evaluated.

The first publication (Zollinger, Philips, and Kumza, 1984) was rated under the mediated model in the category A (ns), indicating that the evidence from the publication did not support the hypothesis. The second publication (Philips, Kumza, Beeson, and Lotz, 1986) was rated under the mediated model in the category A (positive), denoting that the evidence for the hypothesis was conclusive and positive. In the final category, which Koenig, McCullough, and Larson (2001) considered to be of excellent quality, one study met the standards to be evaluated by A-Levels-of-Evidence strategy. The study performed by Enstrom (1989), under the mediated model and independent model, was rated in the category of A (positive), denoting that the relationship between denomination and cancer risk and mortality remains regardless of the introduction of demographic characteristics or selected health practices.

It can be suggested that after examining the ratings provided by the systematic review on all subsections there were enough high-quality publications to infer a conclusion based on the data from the studies. In relation to the mediated model there were three supportive A studies, implying that the evidence was considered to be *persuasive*. Therefore, there is support for the hypothesis that religious denomination is a factor protective of cancer risk and mortality. In terms of the independent model, there were a mixture of A and B studies with an overwhelming amount of them suggesting that the evidence did not support the hypothesis. Therefore, the evidence is described as *consistent failures*. This finding could infer that it is the utilisation of health practices, which completely accounts for the observed association between the two variables.

Hypertension

Hypertension is a condition that is characterised by the individual having higher than normal blood pressure. Suffering from high blood pressure places stress on the heart, which can contribute to coronary artery disease, heart attack, or a stroke. It has been estimated that in the United States more than 50 million people (1 in 4 adults) have the condition (American Heart Association, 2001). Similarly, it has been stated that within the UK more than 41% of men and 33% of women suffer from the condition (British Heart Foundation, 2002). With such high numbers of individuals suffering from this chronic illness makes it an important topic to study, as well as, understanding the mechanisms that protect against this fatal condition. One variable that has received a growing amount of attention over recent years is religiousness.

Levin and Vanderpool (1989) claim the literature reports on a number of studies, which indicate that there is an association between measures of religion and blood pressure. In addition, epidemiologists are aware that significant associations exist between these two variables. These epidemiologists claim that many clinicians are aware of the anecdotal evidence regarding patients whose high blood pressure was helped or better managed because of some regimen or some system of meditation sanctioned by a religious group. The general conclusion surrounding the association is that religion plays a salutary role in the reduction of high blood pressure. One of the first reviews examining the effects of religion on physical health reported that roughly 19% of the studies appraised included, analyses of the effects of religion on cardiovascular disease, and many in fact focused specifically on hypertension (Levin and Schiller, 1987).

Levin and Schiller (1987) report that several studies concluded that certain religious populations appeared to be at higher risk of hypertension. For example, Yemenite Jews with respect to Gentiles (Toor, Agmon, and Allalouf, 1954) are at higher risk. Protestant females pertaining to non-Protestant (Catholics and Jews) females (Ross and Thomas, 1967), Seventh-day Adventists with respect to non-Adventists (Armstrong and Van Merwyck, 1977), and Zen Buddhist Priests in favour of laypeople (Ogata, Ikeda, and Kuratsune, 1984). In addition, the less religious with respect to the highly religious are at greater risk from developing hypertension (Levin and Markides, 1985). A later review by Levin and Vanderpool (1989) elaborated on the hypertension section reviewed by Levin and Schiller (1987). These reviewers augmented the literature by incorporating several more recent studies performed at the time. Similarly, Levin and Vanderpool (1989) reported that in relation to religious affiliation, certain religious groups were found to be at significantly lower risk from suffering hypertension. For instance, epidemiologic studies examining Seventh-day Adventists (e.g., Rouse, Armstrong, and Beilin, 1982; Webster and Rawson, 1979) and Mormons (e.g., Lyon, Wetzler, Gardner, Klauber, and Williams, 1978) would report that members from these populations are at lower risk of suffering from high blood pressure. Levin and Vanderpool (1989; p.76) conclude by saying that “nearly 20 empirical studies reviewed suggest that ‘characteristics and functions of religion have salutary effects on blood pressure’.

Koenig, McCullough, and Larson (2001) commented on numerous researchers that have investigated this condition in association with religion. From the review it can be observed that 36 studies assessed the relationship between the two variables, with the majority reporting that religiousness had salutary relationships with hypertension.

From the catalogue, it was assessed that few studies had utilised denomination as a measure of religiousness. It can be observed that nine studies (29%) employed a measure of denomination in its analysis. From the nine studies, none were rated as weak; similarly, just three (27%) of the studies received the average rating of six.

Average Publications (Rating 5-6)

The three studies from this subsection, according to the reviewers, reveal that certain groups are more likely to suffer from hypertension in comparison with other groups. From the research in this subsection, it was reported that Seventh-day Adventists were less likely to have high serum cholesterol levels in comparison to non-Seventh-day Adventists in Norway (Fonnebo, 1992). Mormons had higher blood pressure than Seventh-day Adventists did in west Australia (Rouse, Armstrong, and Beilin, 1982) and there were no differences between Jews, Catholic, and Protestants in hypertension levels (Ross and Thomas, 1965). The study reporting no differences between the groups employed a sample of medical students thus potentially biasing the results. It can be proposed that medical students are typically of a young age and since hypertension generally affects individuals of an older disposition (Epstein and Eckoff, 1980), employing medical students as a sample may not yield differential effects.

The studies reporting positive associations between the groups utilised cross-sectional designs and employed only low-order analysis, incorporating some controls. For instance, Fonnebo (1992) assessed the coronary risk factors in Norwegian Seventh-day Adventists. The study consisted of 247 Seventh-day Adventists and controls. Results showed that in relation to standardised mortality ratios using the general population as a reference group, Seventh-day Adventists men had a significantly low

standard mortality ratio of 82 ($p < .001$), while the standard mortality ratio for female Seventh-day Adventists was 95 (ns). However, assessing the differences between Seventh-day Adventists and controls in relation to blood pressure levels it was reported that there were no significant differences between the two groups in men. However, both systolic and diastolic pressure were lower in Seventh-day Adventist women compared with controls ($M = 120.8$ vs. $M = 123.4$, $p < .05$ for systolic BP and $M = 75.5$ vs. $M = 77.8$, $p < .05$ for diastolic BP). However, since the primary aim of this study was to assess coronary risk factors, the researchers employed hypertension as a risk factor rather than a primary variable of investigation. In addition, this was a case-control study in which the subjects were matched in relation to age, sex, and study area. The strategy of matching ensures equivalence of groups only on the matched variables. As a result, the groups could differ with respect to health behaviours; the analysis did not control for these factors.

The evidence provided by these publications typically report that denominations, particularly those, which are influenced by strict doctrinal practices, report lower blood pressure levels than denominations, which are more liberal. However, the studies that represent this subsection contain the usual methodological flaws inherent in much research surrounding the area between religion and health, for example the publications come from cross-sectional studies and they do not incorporate the potential influence of many intervening variables, and confounders within their analyses. Therefore, in terms of evaluating the evidence under A-Levels-of-Evidence strategy, all the empirical research within this subsection, would receive a rating of C, which means that the evidence is inconclusive.

Good Publications (Rating 7-8)

The majority of the studies (56%), within this subsection, were rated as good quality. Therefore, the publications received a ranking of seven to eight by the reviewers. The results from the five studies in this subsection generally reported similar conclusions to the previous sections; individuals from certain religious denominations have lower hypertension than those from other denominations. For instance, Seventh-day Adventists in Australia (Armstrong, Van Merwyk, and Coates, 1977; Webster and Rawson, 1979) and those affiliated with a religion (Livingston, Levine, and Moore, 1991; Stavig, Igra, and Leonard, 1984). However, Brown and Gary (1994) reported that within a community sample of 537 African-American males, hypertension was unrelated to religious affiliation. However, the researchers found that those respondents who indicated that they belonged to a denomination were less likely to smoke and consume alcohol compared to the non-denominational group. However, it was observed that religious affiliation was related to less smoking behaviour ($p < .001$) and less daily drinking ($p < .05$).

The significant difference between the studies from the previous subsections and the studies reporting a positive association within this present subsection is the vast majority of the research, within this subsection, incorporated multiple controls in the analysis. For instance, Livingston, Levine, and Moore (1991) examined the social integration and Black intraracial variation in blood pressure, which incorporated a religious variable, affiliation. From the analysis, controlling for multiple covariates, it was reported that affiliation was significantly related to lower systolic ($\beta = -4.90$, $p < .05$) and lower diastolic ($\beta = -6.51$, $p < .001$) blood pressures in men. Among women,

church affiliation was significantly related diastolic ($\beta = -5.32, p < .001$) blood pressure.

A characteristic of the studies in this subsection is that they utilised cross-sectional designs, and the limitation with this type of design is that religious involvement and physical health status are measured simultaneously prohibiting decisions on the likelihood of cause or consequence. Therefore, in terms of the systematic review process, the evidence generated by these studies would not meet the methodological standards to be evaluated; hence, they receive a rating of C (inconclusive).

Excellent Publications (9-10)

One study, within this subsection, received a rating of nine to ten, thus indicating that the researcher conducted an empirical study that was considered very good or excellent, by the reviewers. The distinction with this study is that it employed multiple controls and utilised a longitudinal design. The study performed by Timio, Verdecchia, Venanzi, Gentili, Ronconi, Francucci, Montanari, and Bichisao (1988) investigated the effects of age on blood pressure changes over twenty years in nuns in a secluded order in Italy. The prospective study consisted of 144 white nuns belonging to a secluded monastic order and 138 white control laywomen. The participants were followed for twenty years to investigate whether living in a stress-free environment influences the effects of aging on blood pressure.

The 144 participants were selected as they represented a human model of life differing in many aspects from the standard in westernised societies. The differential characteristics found within this group in comparison to the control group was in their

levels of anxiety for their future, drive, competition, human expectations, economic and familial stress, social and political tensions, and exposure to noise and pollution.

The survey included a clinical examination, anthropometric and blood pressure measurements, humoral investigations, and questionnaires. Questions concerning demographics, familial and personal medical history, level of education, and personal hygiene (tobacco, alcohol consumption, oral contraceptive use, and dietary pattern) were noted. Blood samples were drawn for determination of cholesterol and triglycerides; in addition, a random sodium measurement was made in a group of 72 nuns and 67 control laywomen. Baseline data was taken in 1994 and follow-up data continued for 20 years (year 4, 8, 12, 16, and 20) for the two groups.

Results from the baseline data showed that there were no significant difference between the two groups on demographic, anthropometric, clinical, and laboratory data, with the exception of serum cholesterol and triglycerides levels, which were higher in the nuns than in the control group (both $p < .001$). Throughout the 20-year follow-up, it was recorded that the two groups showed a highly significant difference ($p < .001$) in both systolic and diastolic values. Regression equations of systolic and diastolic blood pressure on age of nuns and controls by age at study entry (21-30, 31-40, 41-50) showed that the slope of the systolic, as well as the diastolic blood pressure increased with age, and was significantly higher in the controls than in the nuns. In the nuns, the slope typically approximated the zero level.

The researchers concluded that at baseline the two groups did not differ with respect to blood pressure levels, age, race, ethnic background, age at menarche, family history

of hypertension, and body mass index. Therefore, the main factor differentiating the two groups of women seemed to be lifestyle. That is, the secluded niche of the nunneries is virtually devoid of conflict, aggression, and competition for power and money. Living for 20 years in such a low-stress environment appeared to prevent the increase of blood pressure with age in for the sample of nuns. These results are comparable in research observing some unacculturated societies, such as Melanesian Gaus of Fiji (Maddocks, 1967), natives of New Guinea (Whyte, 1958), and Brazilian Indian Tribes (Lowenstein, 1961). Results were not affected by lifestyle behaviours such as smoking, alcohol consumption, tea and coffee consumption, and oral contraceptive use as these products were absent or equally distributed between the groups. Therefore, the basic mechanisms of this phenomenon remained unknown, and the research conducted by Timio, Verdecchia, Venanzi, Gentili, Ronconni, Francucci, Montanari, and Bichisao (1988) did not employ any psychosocial variables to substantiate the proposition, which they concluded with.

The publication within this subsection was a longitudinal study that spanned over 20 years examining the differences between nuns in a secluded order and laywomen controls. In terms of the systematic review process, the study would not be considered adequate standard to be evaluated, with the reasoning that the religious variable was not adequately measured. The population that was assessed was nuns in a secluded compared to laywomen controls, no measurement of religious denomination was referenced.

Koenig, McCullough, and Larson (2001) cited nine publications that assessed the relationship between religious denomination and hypertension. However, from these

nine studies only one of the publications appeared to meet the minimal methodological standards to be evaluated under the systematic review strategy. However, the study performed by Timio, Verdecchia, Venanzi, Gentili, Ronconi, Francucci, Montanari, and Bichisao (1988) did not examine religious denomination per se, but examined secluded nuns compared to laywomen. Consequently, it is not possible to evaluate the study within the realm of religious denomination, which signifies that there was not enough evidence provided for support of the hypothesis, which states that religious denomination protects against hypertension, to infer definite conclusions. Subsequently, the published article is rated in the C category.

Assessing the published articles in all the subsections looking at this domain, religious denomination and hypertension, it first appears that the relationship is a positive one. Certain denominations, particularly those who practise strict religious doctrines, suffer less levels of hypertension. This influence of religious lifestyles can be validated when health practices were controlled in the analyses, as the strength and the significance level of the association dropped substantially. Examining, the hypothesis, that certain religious denomination protects against hypertension, utilising a objective and systematic process, it was viewed that none of the studies published were considered to meet the minimal methodological standards to allow for the evaluation of the hypothesis. Suggesting the need for larger amounts of high-quality publications.

Heart Disease

Heart disease is a chronic illness that affects a great number of people around the world. Heart disease is a condition that is particularly prevalent in the UK and USA.

The main causes of heart disease are lifestyle factors such as smoking, alcohol use, diet, and lack of exercise. Consequently, it is prevalent in cultures, which participate in these behaviours.

Early studies assessing the relationship between health, both mental and physical, typically examined the differences across religious groups. In relation to the early literature examining heart disease, the findings tend to be mixed, with empirical research reporting that certain religious groups having higher prevalence rates of heart disease (e.g., Epstein and Boas, 1955; Epstein, Simpson, and Boas, 1956). However, some researchers have reported that there was no difference between religious groups (e.g., Skyring, Modan, Crocetti, and Hammerstrom, 1963; Winkelstein and ReKate, 1969; Winkelstein, Stenchever, and Lilienfeld, 1958). Studies that report a positive association generally find a relationship between those from a Jewish tradition and coronary artery disease. For example, Friedman and Hellerstein (1968) examined a sample of 2,342 attorneys from the Cleveland and Detroit area and reported that in the 40-59 age bracket, Jewish men had significantly more coronary artery disease than non-Jews (13.2% vs. 3.3%, $p < .01$). However, these findings were uncontrolled.

Research that showed no association generally examined the differences between those in mainline Christian denominations (Catholics vs. Protestants). For instance, Brown and Ritzmann (1967) compared 133 patients (65+) who were free of heart disease and 100 control patients (65+) who were known to suffer from this condition. Among the disease free patients 61% were Protestants, 15% were Catholic, Jewish, Greek Orthodox, or Islamic, and 24% had no religious preference. Among the controls, 47% were Protestants, 31% were Catholic or other religious faiths, and 22%

had no religious preference. The researchers concluded from their study that there was no association between coronary artery disease and religious affiliation. In a more recent study, Watson (1991) reported that the higher the percentage of Catholics in the Western European countries under investigation, the lower the rate of death from coronary artery disease. Watson (1991) studied the percentage of Catholics in 24 "Christian" Western European countries. He correlated these percentages with coronary artery disease mortality rates per 100,000 for each country. The correlation between the percentage of the population that was Catholic and the coronary artery disease mortality was $r = -.59$ ($p < .001$). However, control variables, such as socio-economic status were not taken into account.

Throughout the past few decades, the consequence that risk behaviours have on chronic illnesses have been increasingly recognised. Investigators realised the importance of participating in "good" health behaviours (i.e., alcohol avoidance, no smoking, and exercise adherence) in protecting against conditions, such as coronary artery disease. Consequently, over recent years researchers have begun to turn their attention to examining religious affiliates who practise these good health behaviours, such as Mormons and Seventh-day Adventists. Enstrom (1989) performed a prospective study involving 9,844 (53% were high priests and 47% were wives of high priests) religiously active Mormons who lived in California. The Mormon group was compared against a control population of 3,119 non-Mormon adults. Results showed that Mormons had fewer cardiovascular diseases and lower mortality rates in comparison with non-Mormons. Similarly, Philips, Kuzma, Beeson, and Lotz (1980) reported the results of a 16-year prospective cohort of 22,940 California Seventh-day Adventists and a 12-year prospective study of 112,726 California non-Seventh-day

Adventists. All subjects were white and over the age of 33. Results from the study showed that rates of coronary artery disease mortality was 34% lower in Seventh-day Adventists men compared to non-Mormon controls; however, there were no significant differences between the female groups. Conversely, Frazer, Strahan, Sabate, Beeson, and Kissinger (1992) examined coronary artery disease risk factors within a community involving 27,658 Seventh-day Adventists for 6 years. Findings revealed that Seventh-day Adventists, in this population, had a lower risk of coronary artery disease compared to other population groups. In addition, the researchers reported that lack of exercise, obesity, and cigarette smoking were significantly related to coronary events. This led the investigators to conclude that the epidemiology of ischemic heart disease in Seventh-day Adventists was qualitatively similar to that seen in other population groups who practiced these “good” health behaviours.

These findings have led researchers to conclude that Mormons and Seventh-Day Adventists experience substantially less heart disease in comparison with individuals from the general population. Belonging to these denominations, generally result in a healthier diet and more positive health behaviours. These positive health behaviours are largely based on the religious beliefs and commitments practiced by individuals from the religious populations.

Koenig, McCullough, and Larson (2001) reported on thirty-one studies that assessed the relationship between religiousness and heart disease. A large percentage of the studies, 70%, employed religious denomination as a measurement item in the analysis. The findings frequently reported from these studies were that certain

denominations have a greater risk of heart disease problems compared with other denominations.

Weak Publications (Rating 1-4)

The percentage of studies that received a rating of one to four in this category was minimal. According to the reviewers, four studies were regarded as weak in their sampling method, quality of religious measure, quality of statistical analysis, interpretation of results, and the discussion in the context of the existing literature. However, the findings reported from the studies revealed that there is generally no association between religious denomination and heart disease. This lack of an association was found in a case-control study between Catholic and Protestant denominations (Skyring, Modan, Crocetti, and Hammerstrom, 1963) a case-control study between Catholics, Jews, and Protestant traditions (Winkelstein and Rekate, 1969), and those affiliated with a religion in comparison to those not affiliated (Brown and Ritzmann, 1967). However, one study in this category revealed that there were differences between denominations in relation to myocardial infarction (Wardwell, Bahnson, and Caron, 1963), after controlling for age, sex, and race.

Wardwell, Bahnson, and Caron (1963) conducted a study examining the social and psychological factors in coronary heart disease. The case-control study consisted of 32 white male survivors of myocardial infarction, ranging in age from 35 to 64, and 32 age-matched white male controls free from coronary heart disease. The participants were selected from clinical evaluations and from hospital electrocardiographic records and were drawn from the same geographical area.

The participants were interviewed, and asked to fill out six checklists scales, three trait-rating tasks, a sentence-completion task, and scales measuring “self-esteem”, “impulsivity”, and a scale measuring “occupational pressure”. In order to determine personality factors (apart from sociological and demographic variables) the researchers matched 24 pairs of coronary patients and sick controls on the following variables; age (within 5 years), occupational level, religion (in the same major grouping), and ethnic background.

Findings from the study revealed that one powerful variable that influences myocardial infarction was religious affiliation. Protestants showed a myocardial infarction rate nearly four times as great as Catholics. However, the study employed religious affiliation as one of numerous background variables, whilst not controlling for the effects of other demographic variables. In addition, the research was conducted at a time when little work was being carried out, which examined the potential effects of psychosocial factors on specific chronic illnesses. Therefore, the consequence of lifestyle factors were not considered as possible moderating factors influencing the relationship between religion and heart disease. In addition, the study suffered from problems that could affect the non-representativeness and produce sampling errors, since the sample consisted of 64 participants for both groups and the data only consisted of one year’s experience in one country, the findings from the study should be regarded as tentative. The study presented a relationship between the variables that is fraught with methodological limitations, which inhibits a valid association being declared.

The majority of the studies within this subsection revealed that religious denomination was not associated with heart disease, in a variety of populations, including elderly individuals, adults over the age of 45, and white elderly female medical patients. Although, one study did report that Protestant medical patients were four times more likely to be associated with heart disease than Catholic medical patients. However, in examining the evidence under the systematic review approach, none of the studies met the minimum standards as to be utilised by the approach. The main reason for the inadequacy is that all the studies employed cross-sectional designs, a methodological problem frequently experienced by research within the domain of religion and health, both mental and physical. Therefore, the rating that granted to the publications within this subsection (under the systematic review process) would be C. Inferring that the evidence provided by the publications are inconclusive.

Average Publications (Rating 5-6)

There were eight studies categorised in the review with average rankings (5-6). The findings from the results generally revealed that there were significant differences between religious denominations. However, the studies within this subsection rarely employed the use of controls in the analysis. The studies reported similar findings to those from previous subsections; members of certain denominations, specifically those outside of strict orthodox doctrine, are at an elevated risk from suffering from chronic illnesses such as heart disease. For instance, differences were found between Ashkenazi Jews and Oriental Jews in Israel (Dreyfuss, 1953). Although, controls for socio-economic were not utilised, and since Oriental Jews come from a lower social class, there were less likely to part-take in behaviour such as smoking (Dreyfuss,

1953). Non-Mormons reported more heart disease than Mormons (Jarvis, 1977) and non-Seventh-day Adventists over Seventh-day Adventist's (Philips, Lemon, Beeson, and Kumza, 1978). Friedman and Hellerstein (1968) and Ross and Thomas (1965) found that Jewish men reported higher rates of heart disease in comparison with non-Jews. In addition, it has been reported that Protestants reported higher rates of heart disease than Catholics (Wardwell, Hyman, and Bahnson, 1964; 1968); yet, Wardwell and Bahnson (1973) in a case-control study reported that Catholic cases stated higher levels of heart disease than Protestant or Jewish cases.

The studies in this subsection reported a positive association between denomination and heart disease, the majority of them utilised zero-order, or low-order analysis. In addition, the studies typically examined medical samples thus reducing the generalisability of the findings. However, one study, investigated coronary heart disease mortality among Seventh-day Adventists with differing dietary habits, employing multiple controls and one of the studies used adults from a community sample.

The general conclusions offered by the publications within this subsection is that certain religious denominations are associated with an elevated risk than other religious denominations, the evidence from the publications can not be utilised under A-Levels-of-Evidence approach, due to the fact that they employed a cross-sectional methodology. Therefore, the rating granted to the publication is C, which means that the evidence is inconclusive.

Good Publications (Rating 7-8)

Few studies within this subsection, investigating denomination and heart disease, received the ranking of good (7-8). Once again, it has been reported that certain dominations, particularly those who practiced optimum health behaviours are at a reduced risk of suffering from heart disease. For instance, within this subsection it has been reported that Orthodox Jews have significantly lower risk from heart disease than Secular Jews, after controlling for multiple controls (Friedlander, Kark, and Stein, 1986), non-Mormons and non-Seventh-day Adventists in comparison to their counterparts, Mormons and Seventh-day Adventists (Lyon, Wetzler, Gardner, Klauber, and Williams (1978). In addition, few researchers found that individuals from a Jewish tradition reported higher rates of heart disease in comparison to Italians respondents (Epstein and Boas, 1955; Epstein, Arbor, Simpson, and Boas, 1957). Although, the studies published present positive findings the majority of them employed few controls in their analysis with the exception of Friedlander, Kark, and Stein (1986) wherein multiple controls were utilised. Friendlander, Kark, and Stein (1986) examined the relationship between religious orthodoxy and myocardial infarction in Jerusalem. The case-control study consisted of 1,225 participants, of which 539 belonged to the myocardial infarction group and 686 belonged to the control group. Within the myocardial infarction group there where 454 men and 85 women. In the control group, there were 295 men and 391 women. Results revealed that secular Jews had a significantly higher risk of myocardial infarction compared to orthodox Jews (OR 4.2, 95% CI 2.6-6.6, for men and OR 7.3, 95% CI 2.3-23.0, for women). This relationship was independent of the other variables in the model (age, ethnicity, education, smoking, physical exercise, and body mass index).

When examining the evidence presented by the publications cited by Koenig, McCullough, and Larson (2001), it would appear that certain denominations are associated with an increase risk of suffering from heart disease. However, under the systematic review process, the evidence from these publications were provided from studies that conducted cross-sectional designs, therefore, it is not possible to infer if heart disease is a cause or consequence of religious denomination. In addition, the majority of the research did not utilise high-level controls, such as health behaviours, within their analysis, prohibiting the ability to cancel out these factors in indicating the sole effects of a group's religious denominational relationship with heart disease. Therefore, the studies within this category are rated a C (inconclusive).

Excellent Publications (Rating 9-10)

Within the final subsection in this domain, there was one study that examined denomination and heart disease which was considered of excellent quality to be placed within the excellent rating (9-10). The prospective cohort study conducted by Goldbourt, Yaari, and Medalie (1993) investigated the factors predictive of long-term coronary heart disease mortality among 10,059 male Israeli civil servants and municipal employees.

The researchers aimed to investigate the factors related to the development of coronary heart disease, hypertension, and diabetes in a multicultural society. The study consisted of 10,059 subjects chosen by stratified sampling procedures, and aimed at obtaining approximately equal numbers of study subjects born in five different geographic-cultural Jewish immigrant groups, as well as Jews born in the

pre-1948 British mandate of Palestine. Consequently, those involved in the study included all Jewish men, aged 40 and over, born in central Europe, eastern European born, those born in southeast Europe, all those born in the pre-1948 borders of British-dominated Palestine, those within the post-1948 borders of independent Israel, those from the mideastern Asian countries as well as all such Jewish male employees born in northern Africa. The subjects consisted of a diverse work force such as teachers, social workers, postmen, prison wardens, dock workers, workers in the railway transport system, cleaning personnel, engineers, scientists, and administrators. Providing an extensive representation of the different occupations and educational/socio-economic levels in the country.

The first set of data was recorded in 1963 and consisted of socio-demographic and behavioural information, and responses on a psychosocial questionnaire. The respondents were interviewed for dietary habits and underwent electrocardiographic, biochemical, blood, and genetic (blood group) testing. The major characteristics of the data were reassessed in 1995 on 95% of the original candidates. Additional information pertaining to a number of sociological questions, including topics, such as religious education and upbringing, concentration camp experience during the holocaust period in Europe, status incongruity, and occupational mobility were collected. In 1968 the last follow-up medical and electrocardiographic evaluation of the cohort took place, in which an assessment of the spirometric function was added. At the 23-year follow-up the researchers investigated the factors which predisposed the individuals from coronary artery disease. During this period, 3,473 deaths were recorded including 1,098 for which coronary artery disease was assessed as the underlying cause.

The results from the analysis showed that the highly Orthodox religious groups had the lowest rate of mortality from heart disease in comparison with non-believers (36 vs. 61 per 10,000). However, there were no differences in cases between Orthodox, Traditional, or Secular Jews. However, the variable, religion, only made a serendipitous guest appearance as a background variable, therefore, it was not included in any high-order analysis. From the findings it would appear that the factors predictive of long-term coronary artery disease mortality were age, systolic blood pressure, smoking, cholesterol, high-density cholesterol levels, diabetes, history of myocardial infarction, and angina pectoris, all significant ($p < .001$). It can be inferred that the most orthodox religious groups have optimum health practices that protected against coronary heart disease. However, this was not tested in the analyses and any definite conclusions cannot be stated and validated. In addition, there were no controls for social support or depression; therefore, it is not clear whether this observed association was mediated by known cardiovascular risk factors. In light of these problems, the research performed by Goldbourt, Yaari, and Medalie (1993) does meet the methodological standards to be evaluated by A-Levels-of-evidence approach. The research does not get entered under the mediated model, as there was no control for socio-demographic characteristics. Under the independent model, the publication gets a rating of B positive, which means that the evidence is generally sound showing a positive association between the two variables.

Observing the area which examined the relationship between religious denomination and heart disease, Koenig, McCullough, and Larson (2001) cited 21 studies. The reviewers concluded that

‘While studies on this topic are few, degree of religious involvement also appears to be an important influence of on the risk of myocardial infarction, mortality from coronary artery disease, and survival following coronary artery bypass surgery’ (p. 249).

In examining the published articles cited by Koenig, McCullough, and Larson (2001), it appears that there is a positive relationship between members of certain religious denominations and heart disease. However, the majority of the research within this section did not incorporate controls in their analyses; therefore, not permitting the ability to exclude these confounders as possible variables which may explain the observed relationship. The latter study in this section did control for known risk and protective factors, and the study was rated in the B (positive) category. Not collecting data on socio-demographic characteristics prevents the study being rated in the A category. Subsequently, the study is considered to be generally sound; although, contains some flaws.

In examining the hypothesis that states religious denomination protects against heart disease, under A-Levels-of-Evidence approach, it can be viewed that within the subsections, one of the studies meets the minimal standards to be evaluated. Subsequently, the accumulated research reaches the level of *insufficient evidence* for the hypothesis. This suggests the need for a greater amount of research to be concluded, which is of an adequate standard.

All-cause mortality

For more than a century it has been recognised that certain religious groups have higher rates of morbidity and mortality in comparison to other groups. These observations first led John Shaw Billings (1891) to consider religious affiliation to be

a potential factor in the study of differential rates of morbidity and mortality among social groups. Subsequently, over the years there has been an increase in the numbers of investigations examining the association between religious denominations and all-cause mortality. One of the first reviews assessing physical health was conducted by Levin and Schiller (1987), these authors reported that research is mixed with some writers arguing from the perspective of “holistic health” that religious affiliation is promotive of health (e.g., Berkel, 1983; Dysinger, 1963). Other scholars have noted that religious behaviour can also be a risk factor for stress-related diseases, in that certain religious traditions may create or lead to high levels of guilt or anxiety (Kaplan, 1976; MacDonald and Lockett, 1983).

The section on all-cause mortality includes all types of illness with the exception of cancer, heart disease, and hypertension. The studies typically contrast members and non-members of particular faiths across several causes of death.

Weak Publications (Rating 1-4)

In the review by Koenig, McCullough, and Larson (2001), there were 21 studies cited that investigated the association between religious denomination and all-cause mortality. From these twenty-one studies, only one study was considered weak and received a rating of two. This study conducted by Saksena and Srivastava (1980) compared two religious groups, Muslim and Hindu, on the biosocial correlates of perinatal mortality in an Indian hospital. The study was cross-sectional and consisted of 5,506 births in Lucknow, India, during the period September 1976 to August 1977. The sample obtained for the study came from the largest government-run hospital in

the area, therefore, it was postulated that the study hospital attracts individuals from all socio-economic levels, creating a sample, which is likely to be representative.

The data collected for the study contained information on biological, clinical, and socio-economic status of the infants and their families. The biological data included information on age of mother, parity, period of gestation, birth weight, and sex of the child. Clinical data included information on presentation and type of delivery. Information on religion, family income, and occupation of father formed the socio-economic data file.

In relation to the variable measuring religion, results indicate that perinatal mortality among Muslims was 95.2 per 1000 compared to 72.6 per 1000 for Hindus, a statistically significant association ($\chi^2 = 4.88$, $df = 1$, $p < .05$). The χ^2 statistic was close to zero showing small differences between the groups. In addition, the χ^2 statistic is a low-order level of analyses, hence signaling no controls were utilised. Other factors involved in perinatal mortality included; age of mother, parity, period of gestation, birth weight, and sex of the child (biological factors); and family income and occupation of the father (socioeconomic factors). It can be suggested that low birth weight and period of gestation is dependent on socioeconomic factors, that there is an interaction between these variables that lead to perinatal mortality. In addition, it can be postulated that in terms of religious denomination, the differences in perinatal mortality is potentially the consequence of ethnic and cultural differences between Muslims and Hindus living in India (Miah, 1993). However, the analysis of income and occupation for each religious tradition were not employed in the study prohibiting any suppositions being inferred. In terms of A-Levels-of-Evidence approach, the

evidence from this publication is not deemed to come from a methodology that is of adequate standard to be evaluated, since the publication is cross-sectional and does not employ the use of controls in the analysis. Due to these flaws the study is rated as C, which denotes that the design of the study makes it impossible to rule out bias, confounding, or chance as alternative explanations for results, hence the study is eliminated from further consideration.

Average Publications (Rating 5-6)

The proceeding subsection within this domain, examined the relationship between religious denomination and all-cause mortality for the publications that received the quality rating of five or six. The total number of studies within this subsection equals ten. The main conclusion stated by these studies, are that certain religious denominations have greater likelihood of mortality in contrast to other groups. For instance, it has been reported in large community studies that Seventh-day Adventists men and women generally live longer than men and women from the general population (Berkel and de Waard, 1983; Dysinger, Lemon, Crenshaw, and Walden, 1963; Fonnebo, 1992, 1994), Mormons have lower standardized mortality ratio levels than non-Mormons (Jarvis, 1977), there is less infant/child mortality in Muslims compared to Hindus (Maih, 1993), and Jewish individuals are more likely to die from certain illnesses, e.g., diabetes, than from other illnesses, e.g., Tuberculosis, and those individuals live longer than individuals from the general population (Bolduan and Weiner, 1933; Liberson, 1956; Needleman, 1988; Seidman, Garfinkel, and Craig, 1964). It has been postulated by the researchers that the differences between the denominations are the consequence of lifestyle behaviours (Dysinger, Lemon, Crenshaw, and Walden, 1963; Jarvis, 1977) and socio-economic differences leading to better health care facilities and medical care (Liberson, 1956; Miah, 1993).

The suppositions that lifestyle factors help explain the relationship between religious denomination and mortality is supported by examining the research by Berkel and de Waard (1993). These researchers examined the mortality pattern and life expectancy of Seventh-day Adventists in the Netherlands. The study consisted of 3,217 Seventh-day Adventists (1,066 males, 2,151 females) as the index population and the mortality rates in 1972 of the Dutch population as the reference population. The information on mortality pattern and life expectancy was obtained by calculating the total person-years at risk of the Seventh-day Adventist population (by age) with the corresponding 1972 death rates of the total Dutch population (by age). The observed/expected ratio was then found by dividing the observed deaths by the expected number of deaths. The significance level of the results were calculated according to Bailar (1964). The last part of the study consisted of a "health survey" between a sample of Seventh-day Adventist (n= 190) and a group of friend controls (n= 170). The data collected contained information on medical history, a (short) physical examination, laboratory investigations, and a detailed dietary survey.

Findings from the study revealed that 522 Seventh-day Adventists died during the study period (1968-1977). It was calculated that the life expectancy of Seventh-day Adventists men and women were greater than men and women from the Dutch population (52.3 years and 52.5 years vs. 43.4 years and 47.9 years, respectively). The pattern of mortality reveals differences between the expected and observed deaths in Seventh-day Adventists. The differences in the observed/expected ratio were highly significant. The significant differences were found in total mortality ($p < .01$), neoplasms ($p < .01$), a host of cancers ($p < .01$), cardiovascular diseases ($p < .01$), heart

disease ($p < .01$), and respiratory disease ($p < .01$). In an attempt to discover and explain these differences, the investigators performed a "health survey" among a sample of Seventh-day Adventists and a group of "friend" controls. Findings from the "health survey" disclosed that there were no differences in hypertension levels, the occurrence of diabetes, and physical activity, whereas, differences were found in obesity (controls were less obese), cholesterol levels, cigarette smoking (all lower in Seventh-day Adventist), and the level of fatty acids (higher in Seventh-day Adventists).

The results from the study validated the general conclusions that certain denominations were at less risk from all-cause mortality compared with other denominations. However, the findings from the comparison of known and possible risk indicators between the two groups demonstrated the effect of life-style characteristics, such as abstinence from the use of tobacco and alcoholic beverages and the church's recommendation for a vegetarian diet. Therefore, it can be suggested that if the differences in mortality pattern between the two groups were investigated and these variables were employed as controls then any statistical significant result may potentially be reduced to non-significance. Nevertheless, this study disclosed how a unique lifestyle (non-smoking, more appropriate diet) increases life expectancy.

Initial examination of the publications within this subsection would suggest that certain religious denominations, particularly those that are not influenced by strict religious doctrines, are associated with higher ratio of all-cause mortality. However, in closer examination of the studies it can be viewed that the majority of the studies

come from cross-sectional designs and do not incorporate the utility of high-level controls, but instead utilizes low-order controls, such as age and sex. Although, these extraneous variables help explain some of the observed relationship between religious denomination and health, it has been shown in a vast amount of research that a variety of socio-demographic characteristics are important mediators in the relationship; therefore, they deserve to be included within any analysis between the two variables. Due to these flaws, the studies would be categorized under the strategy of the systematic review as C (inconclusive) and eliminated from further investigation.

Good Publications (Rating 7-8).

The general conclusion within each of these subsections is that certain religious denominations have greater life expectancies than individuals from other denominations. In this subsection, the general conclusion is once again stated. The studies that will be examined, have received a rating of seven to eight by Koenig, McCullough, and Larson (2001). This rating indicates that the studies are considered to be of a very good standard. The conclusions rendered by the researchers are analogous, that is all the researchers reported that certain religious denominations have higher mortality rates and shorter life expectancies in comparison with other denominations. For instance, was reported that perinatal and infant death was less among the Amish in Ohio (Acheson, 1994) and higher among Faith Assembly members in Indiana (Spence, Danielson, and Kaunitz, 1984). Jewish and Amish men live on average longer than their counterparts (Goldstein, 1996; Hammen, Barancik, and Lilienfeld, 1981). Those individuals affiliated with Eastern Orthodox traditions had higher mortality rates than those from Lutheran traditions (Rasanen, Kauhanen, Lakka, Kaplan, and Salonen, 1996). Christian Scientists and non-Seventh-day

Adventists had higher mortality rates than their comparable groups (Simpson, 1989; Webster and Rawson, 1979). Finally, those elderly individuals unaffiliated with a religion had higher mortality than those affiliated with a religion (Zuckerman, Kasl, and Ostfeld, 1984). Although, the studies generally stated differences between one religion and another in terms of mortality patterns and life expectancies, the studies typically utilized controls which were low-order, such as sex, age, and ethnicity. Consequently, it can be postulated that the introduction of higher-level covariates such as health behaviours may provide different conclusions. Subsequently, one study within this subsection adjusted for numerous covariates and found that the results remained consistent (Rasanen, Kauhanen, Lakka, Kaplen, and Salonen, 1996).

The results from the majority of the published articles within this subsection came from research that was cross-sectional and with the majority of the studies utilized controls such as age or sex, without the employment of higher-order controls, such as socio-economic status, health practices, and mental health variables. Consequently, the majority of published studies in this subsection are not of adequate standard to provide enough evidence to establish if there is enough support to suggest a relationship between the two variables. Therefore, the studies are rated as C (inconclusive). However, there were a few published articles, which did not contain the flaws to be excluded from the evaluation process.

The first publication performed by Zuckerman, Kasl, and Ostfeld (1984) was a two year prospective cohort investigation into the psychosocial predictors of mortality among 225 elderly individuals forced to move from their homes and 173 elderly persons (from the same neighbourhood, not forced to move) matched on age, race,

sex, and marital status. The variables of interest to the researchers were religion, well-being, and social contacts. A variety of variables were measured, such as socio-demographics, religion, and health information. No statistical analysis was performed on the variable religious denomination. Consequently, the publication can not be utilized for the systematic process to offer evidence for the hypothesis, religious denomination protects against all-cause mortality and receives the rating of C.

The second study in this subsection performed by Rasanen, Kauhanen, Lakka, and Salonen (1996) was a six-year prospective cohort study investigating the association between religious affiliation and all-cause mortality among middle-aged men in eastern Finland. The study consisted of 2,682 men aged 42, 48, 54, and 60 at baseline, subjects with prevalent coronary heart disease or cancer were deleted from the study (n= 1042). Consequently, leaving 1,624 men with data on religious affiliation for analysis. Analysis from the 1,624 men in the study revealed that the majority of the sample was Lutheran (84%), 10.6% non-affiliated, and 5.4% Orthodox. Results from the analysis revealed that in comparison with Lutheran men, Orthodox men had considerably higher all-cause mortality, even with the inclusion of a variety of confounders (RH 5.12, CI 1.98-13.27, $p < .001$). The findings revealed that there were significant differences in smoking and alcohol use among the religious groups, these did not account for the mortality risk among the religious groups.

The study by Rasanen, Kauhanen, Lakka, and Salonen (1996) demonstrates how religious denomination is associated with mortality, even when a variety of confounders and known risk factors are controlled for in the analyses. In relation to the systematic review, the study did not contain any of the flaws that would omit the

study being evaluated. Therefore, the study can be evaluated under the mediated and the independent model. In relation to the both models, the study is rated as A (positive). This rating denotes that the evidence from the research is conclusive and offers support for the hypothesis.

Excellent Publications (Rating 9-10)

Following from the tradition of the previous subsections, only one study was deemed of a high enough standard to receive a rating in the excellent category. This study rated ten by Koenig, McCullough, and Larson (2001), was a longitudinal study, investigating the mortality rates between members from a religious kibbutzim in comparison to members from a secular kibbutzim in Israel. The 16-year (1970-1985) study conducted by Kark, Shemi, Friedlander, Martin, Manor, and Blondheim (1996) examined the differences between 3,900 subjects from 11 religious kibbutzim and 11 secular kibbutzim.

The study was performed in an aim to assess the association between Jewish religious observance and mortality among study populations with maximal similarity in social structure, social support mechanisms, and lifestyles. Therefore, the hypothesis set stated that belonging to a religious community provides additional protection against mortality. The study matched secular and religious Kibbutz's according to location, use of the same regional hospital, and members older than 40 years, all-cause mortality was significantly greater among members of the secular Kibbutz. This finding would lead to the perception that belonging to a religious organisation is better for an individual's health, but the matching criteria was not adequate and information containing dietary habits, alcohol consumption, and smoking were not

included. The authors reported that the secular group were at greater risk of higher blood cholesterol conditions, but this was not controlled for in the analyses. Inadequate control for important covariates may point to significant findings when none may exist.

It was demonstrated that within this subsection, the relationship between religiousness and mortality is one that is positively linked. However, the foundation for this conclusion is presented from one study, which contained methodological flaws. For instance, in terms of the hypothesis under investigation within this category, that religious denomination is associated with mortality, the variable measuring religious denomination is not adequately measured. Additionally, the study did not control for any confounding variables, such as age or sex. Therefore, the publication under A-Levels-of-Evidence approach is rated as C (inconclusive).

After re-evaluating the studies within this category, it can be observed that the majority of them presented findings that support the supposition that those members from certain religious denomination are linked to lower mortality. However, when the studies incorporated controls into the analyses the associations tend to drop substantially. Additionally, utilising an objective strategy to evaluate the studies it is possible to establish, which of the studies meet the criteria of acceptable methodological standards. From the nineteen studies cited by Koenig, McCullough, and Larson (2001), one of the studies was considered of adequate standard to be evaluated. Due to the number of covariates utilised by Rasanen, Kauhanen, Lakka, and Salonen (1996), the study can be evaluated under models, the mediated and the independent. For the two models the study received the rating of A (positive)

indicating that the study provided conclusive support for the hypothesis, with the relationship not been accounted for by known risk or protective factors. Due to this rating the study demonstrates that the publication provides *some evidence* for the hypothesis, religious denomination protects against mortality.

Conclusion

The suggestion made by the majority of researchers within this domain examining religious denomination and physical health is that there are positive associations between the two variables. Levin and Schiller (1987) comment on how the consideration of religious factors has been a major focus of epidemiologic research for nearly two centuries. In their review, they comment on a number of studies that assessed the relationship between religious denomination and a variety of physical health variables, and list a plethora of research articles that present positive findings of one religious group compared to another. Levin and Schiller (1987) stated

‘Over two hundred empirical studies in epidemiology have included religion variables and have obtained positive findings associating religion and health’ (pp. 23-24).

However, the reviewers found criticisms with the empirical studies. With regard to religious affiliation, Levin and Schiller (1987) suggest that as an operational construct, confining individuals in terms of Catholic, Protestant, Jewish, or Gentile for example, may be little more than a proxy for socio-economic status or ethnicity. They also argue that the literature restricts the categorisation of religious groups disregarding the heterogeneous nature of denominations, for example, Jews can be expanded into Reform, Conservative, and Orthodox classifications. These researchers make a further criticism of the utilisation of religious affiliation as measuring tools in that it does not represent a reliable indicator of individual differences in religiosity,

and that maybe more “interiorised” measures of religiosity such as degree of faith, pattern of practice, or theological orientation may prove to be more fruitful indicators.

A distinguishing feature about the published articles within this chapter is that they generally comment on the protective nature of religion for those denominations that practise religious doctrines. For instance, the majority of studies which demonstrated that Mormons and Seventh-day Adventists were at lower risk from physical ill health, did not demonstrate that it was likely that diet, health behaviours, and quality of social and family life accounted for most of this risk. Research that did incorporate vast numbers of covariates, generally revealed associations between the two variables which diminished.

The review by Koenig, McCullough, and Larson (2001) presents a view, which appears to demonstrate the protective nature of religious denomination on the various physical health areas. However, in closer examination of the studies it can be revealed that they suffer much of the same methodological limitations as those assessing religious denomination and mental health. In assessing the cited articles alongside A-Levels-of-Evidence approach, it can be observed that the published articles do not meet the minimal methodological standards to be evaluated by the process.

In examining the hypothesis, religious denomination protects against cancer, it appears that there is enough high-quality research to make a judgement on this conclusion. Six studies from this subsection offered findings from research which were considered of good standard to be evaluated. The findings from the studies examining this domain were mixed with some presenting no association (Kune, Kune,

and Watson, 1992; Toniolo and Kato, 1996; Zollinger, Philips, and Kumza, 1984) whilst the remaining evidence supports the hypothesis (Enstrom, 1989; Jensen, 1983; Philips, Kumza, Beeson, and Lotz, 1986). The findings from the latter three studies reached the level of *persuasive evidence* for the mediated model. Therefore, there is support for the hypothesis that religious denomination is a factor protective of cancer risk and mortality. In terms of the independent model, there were a mixture of A and B studies with an overwhelming amount of them suggesting that the evidence did not support the hypothesis. Therefore, the evidence is described as *consistent failures*. This finding could infer that it is the utilisation of health practices that completely accounts for the observed association between the two variables.

In relation to the cited publication, which examined the relationship between religious denomination and hypertension, it seems as if the association between the two variables is a positive one. However, when variables were controlled for in the analysis, which examined health practices, the significance level and the strength of the association dropped, sometimes to nonsignificance. Utilising the systematic review to rate the studies demonstrated that none of the cited publications examining this domain were considered to be of adequate standard to be rated by A-Levels-of-Evidence approach. Subsequently, the data reached the level of *insufficient evidence*. Suggesting the need for larger amounts of high-quality publications is required.

In examining the relationship between religious denomination and heart disease, Koenig, McCullough, and Larson (2001) cited twenty-one studies that addressed this area. From the twenty-one studies, one of the cited publications was considered to have met the minimal methodological standards to be evaluated by the systematic

review process. The study performed by Goldbourt, Yaari, and Medalie (1993) received the rating of B (positive) and although, the methodology from the study was generally sound, supporting the hypothesis, the research base was too small to generate a potential conclusion about the hypothesis. Subsequently, the accumulated research reaches the level of *insufficient evidence* for the hypothesis. Similarly, this suggests the need for a greater amount of research to be performed, which is of an adequate standard.

When the cited publications were examined in the subsection assessing religious denomination and mortality, initially it appears that the relationship between the two variables were favourable. Certain denominations appear to be linked with decreased mortality. However, it was viewed that research which incorporated a variety of health behaviours into the analysis, observed the strength of the association drop. From the nineteen published articles, cited by Koenig, McCullough, and Larson (2001), one of the studies met the standards to be included in the systematic review process. The study performed by Rasanen, Kauhanen, Lakka, and Salonne (1996), utilised a vast amount of covariates in their analyses, prohibiting the ability to rate the data under the mediated and independent models. For the two models, the study received a rating of A (positive). This rating denoted that the data supported the hypothesis and the methodology of the research did not contain any major flaws. This rating for the study indicated that the publications reach the level of *some evidence* for the hypothesis.

Empirical researchers have investigated the four physical health topics under investigation in this chapter. However, it was viewed that the majority of existing

studies did not meet the minimal methodological standards to be evaluated by an objective and systematic process. The studies which were evaluated, demonstrated that there was evidence for religious denomination influencing the physical health of individuals, in relation to certain chronic conditions. Similarly, Koenig, McCullough, and Larson (2001) comment in their review that the relationship between two variables is one that exists, and although, the reviewers also accept that the relationship is partially mediated by the health practices of those in certain denominations, Koenig, McCullough, and Larson (2001) regard the existence of the relationship with too much optimism. The typical and often criticised features of the studies are the samples and measures are incomparable. The studies often use low-order analyses with a religion variable that was considered as just one of numerous psychosocial or behavioural constructs to be measured.

3 Public Religious Activities and Mental Health

Introduction

The subject of this chapter is the relationship between public religious measures, specifically, frequency of attendance at religious services, and its effects on mental health. Employing frequency of attendance as a measure of public religiousness has become increasingly popular over recent years and the measure is readily included in large surveys. The area examining the relationship between frequent attendance at religious services and mental health is one that is frequently examined by measuring some form of public religious involvement, typically church attendance and depression. The measures usually consist of one-item or two-item scales that are fraught with psychometric limitation. Although in spite of these shortcomings inherent in such measures, most studies remain to utilise them as testing batteries for research purposes. Researchers have shown there to be an inverse relationship with the negative aspects of mental health, such as depression (Koenig and Fritterman, 1995; Koenig, George, and Peterson, 1998; Koenig, Pargament, and Nieslen, 1998) and anxiety (e.g., Tapananya, Nicki, and Jarusawad, 1997). Additionally, it has been reported that there are positive associations with life satisfaction (e.g., Levin, Chatters, and Taylor, 1995; Markides, 1983; Musick, 1996; Willits and Crider, 1988), and well-being (e.g., Krause, 1993; Singh and Williams, 1982).

When researchers employ organisational measures of religiosity, such as attendance at religious services, or membership in religious organisations salutary effects with mental health have been presented, regardless of how minimal the association is.

The review executed by Koenig, McCullough, and Larson (2001) presents an abundance of empirical research that investigates the area; ranked on a scale from one to ten. The intention of this chapter is to outline the studies presented in the review and to document the problems associated with each group. Additionally, studies recorded, as being the principal pieces of empirical research in this area will be assessed under A-Levels-of Evidence approach.

Depression

McCullough and Larson (1999) performed a review that assessed the literature specifically on religion and depression. These reviewers reported that within the 29 studies that examined the cross-sectional association on Church attendance, 24 of the studies found that people involved in a religious organisation to have lower levels of depressive symptoms. Stringent control for a number of demographic, psychosocial, and health-related variables, the association between the organisational religious behaviour variable and depression variable drops substantially.

Weak Publications (Rating 1-4)

Koenig, McCullough, and Larson (2001) referred to forty-nine studies that employed measures of organisational religious activity in relation to depression. From the catalogue, it was observed that a small amount of the studies (6%) within this domain, were ranked as weak. Findings from the three studies revealed that frequency of religious attendance was unrelated to depression (Alvarado, Templer, Bresler, and Thomas-Dobson, 1995; Bishop, Larson, and Wilson, 1987; Farakhan, Lubin, and O'Connor, 1984). For instance, Alvarado, Templer, Bresler, and Thomas-Dobson (1995) asked respondents 'How often do you attend an organised service or church

group of some sort?' as one of a series of religious items and correlated the question with a 17-item measure of death anxiety, the items correlated weakly ($r = -.10$, ns). Bishop, Larson, and Wilson (1987) investigated the religious life of individuals with affective disorders. These researchers performed a case-control study of 64 inpatients with major affective disorders and compared them to 109 unmatched controls with psychiatric disorders. Results revealed that there were no differences in current religious activities between cases and controls. Finally, Farakhan, Lubin, and O'Connor (1984) reported no association between church attendance and depression.

However, the studies cited suffer from methodological limitations; all the studies employed small numbers of convenience samples. For instance, the study conducted by Farakhan, Lubin, and O'Connor (1984) consisted of 30 respondents, 23 females and 7 males, aged between 52 to 97 years ($M = 71.4$, $SD = 7.6$). The respondents were recruited from community centres, private residences, apartment buildings for the aged, hence constitute a convenience sample. In addition, the study did not look at religiousness specifically, it was utilised as a background variable. However, the remaining studies did employ religiousness as a primary predictor. Although the study presented by Alvarado, Templer, Bresler, and Thomas-Dobson (1995) generated almost half of the 200 respondents from undergraduate courses and the study by Bishop, Larson, and Wilson (1987) utilised 173 unmatched case-control psychiatric patients.

In relation to the A-Levels-of-Evidence strategy, the three studies would fall into category C and be eliminated from evaluation as the first two studies utilised a cross-sectional design strategy. Although, the third study employed a longitudinal analysis,

the researchers did not incorporate controls into the analysis and in addition they inadequately measured the religiousness variable (church attendance), prohibiting the study from being further evaluated. Consequently, the evidence presented from these three studies is inconclusive ruling out any potential conclusion being inferred surrounding the association between frequent attendance of religious services and depression from these publications.

Average Publications (Rating 5-6)

Within the higher category (five and six), Koenig, McCullough, and Larson (2001) cited fourteen studies, which examined how the concept of public religious behaviours was associated with depression. The reviewers communicated that conclusions from the research generally stated that the variable offers beneficial outcomes for depression. The vast majority of the studies report negative associations between the two variables; the more individuals report that they attend public religious activities the less depressive symptomology they were likely to declare. For instance, this finding was partially supported by Koenig, Moberg, and Kvale (1988) in a sample of 106 elderly medical patients, ranging in age from 54 to 94 years ($M=74.4$, $SD=7.5$). Religiosity was conceptualised by an 88-item questionnaire, extracted from four instruments other researchers have used to capture the various dimensions of religiosity. In relation to the variable of interest, public religious activities, 2-items were utilised that consisted of, frequency of church attendance, responses were on a five-point scale ranging from 'never' to 'several times per week' and frequency of other group-related activities such as adult Sunday school, Bible study, and prayer group related activities. Presence or absence of depression was utilised as the mental health variable. Results from the analysis demonstrated that frequency of church

attendance was high, over 50% reporting at least weekly attendance. Slightly more than a quarter (27%) of the sample participated in weekly adult Sunday school, prayer, or Bible study groups. Around one fifth (24.1%) of the respondents reported that they suffered from depressive symptoms. Employing Wilcoxon Rank Sum Test in analysing the relationship of organisational religious activity and the presence or absence of depressive symptoms, results revealed that patients who reported depressive symptoms reported lower levels of organisational religious activity; although, this finding was nonsignificant for the whole group. However, when the group was stratified by sex, religious behaviour was lower among women with depressive symptoms than among those without such symptoms. However, this was significant at the $p < .10$ level (this was deemed to be a statistical significant level by the researchers since it was an exploratory study).

Kroll and Sheehan (1989) reported findings in 52 psychiatric patients (35 < 34 years and 17 > 35years). Public religious activities were ascertained from a measure of religious attitudes and involvement. Eight items contained questions, such as 'I am a church member', 'I attend church weekly', and 'I help in church'. These were specifically relevant to public church practices. The mental health variable, depression, was ascertained by the DSM-III administered by the resident and attending psychiatrist. In addition, respondents filled out the Beck Depression Inventory (Beck, 1967). Results from the study were presented as frequencies, i.e., females were more likely to be church members, attend church more frequently, and to help in the church. No statistically analyses were provided, such as parametric or non-parametric tests, which could explain the differences between groups or strengths of relationships.

However, Sherkat and Reed (1992) reported the positive effects that church attendance had on depression for 156 suddenly bereaved adults. Church attendance was measured by a single-item from 'never' (1) to 'often' (4) and depression was measured using an 18-item scale, which indicated the frequency of a number of symptoms associated with depression, such as crying, loss of appetite, and feelings of hopefulness. Results from the ordinary least squared analysis reveals that frequency of church attendance was negatively related to depression ($p < .05$). However, when the indices of social support (i.e., social integration, quality of support, and frequency of confiding in friends) were added to the analysis, frequency of church attendance was no longer significant. This finding could infer that depressive symptoms are reduced by church attendance only insofar as church attendance enhances the social integration, the utilisation of social support, and the quality of expressive social support. Finally, Wright, Frost, and Wisecarver (1993) found that in 451 adolescents greater attendance at religious services was inversely related to depression as measured by the Beck Depression Inventory (Beck, 1967), in an uncontrolled analysis.

Additionally, few studies reported no association between the variables. For instance, Blaine and Crocker (1995) reported no significant associations in a sample of 144 psychology undergraduate students. The respondents were asked about their religious participation by estimating how often they attended religious services in the past month; responses ranged from 0 to 22. Depression was measured by using the 13-item short form of the Beck Depression Inventory (Beck and Beck, 1972). Findings from the zero-order correlations reveal that religious participation was not statistically

associated with depression for either black or white respondents. Comparable findings were reported by Schafer (1997) in a sample of 282 sociology students. Neeleman and Lewis (1994) reported that there were no differences between three groups of psychiatric patients and one group of orthopedic patients (control group) in terms of frequency of church attendance.

In assessing a more specific group, Spendlove, West, and Stanish (1984) reported that zero-order associations were present between church attendance and depression in a sample of 179 white Mormon women. However, this significant association diminished whenever education, caring from spouse, health, and income were controlled. Additionally, this lack of a significant association was also reported in a community sample involving 1,125 Mexican-Americans (Levin and Markides, 1985).

Within this subsection rated as average, one study reported the negative effects of frequent religious attendance on depression for teenage mothers (Sorenson, Grindstaff, and Turner, 1995). These researchers presented a longitudinal study of 261 pregnant adolescents. Subjects were drawn between 1984 and 1986, primarily from caseloads of family physicians and obstetricians/gynaecologists. Depression was measured by the Center for Epidemiological Studies Depression Scale (Radloff, 1977) and religious practices was measured by 1-item asking respondents to indicate frequency of church related activities, with responses ranging from 'never' to 'more than once a week'. Results in relation to church attendance revealed that for all respondents, during pregnancy, individuals who reported that they almost never attended/never attended religious services had the lowest scores on the depression scale ($M= 14.44$), followed by respondents who reported they attended services once

a month or more ($M= 17.12$), and finally by those who reported that they attended church a few times a year ($M= 17.88$). When the group were stratified by marital status, Sorenson, Grindstaff, and Turner (1995) reported that within the married respondents, the group to have the lowest depression scores were the respondents reporting attendance at religious services a few times a year ($M= 7.77$). Amongst the unmarried respondents, the group who reported that lowest depression scores, were those who responded to attending religious services almost never/never ($M= 14.68$). Teenagers who had attended religious activities reported higher levels of depression ($t= 2.80, p= .006$).

The studies within this subsection, which received a rating of five and six by Koenig, McCullough, and Larson (2001), would have been mostly excluded from the systematic review and received a rating of C, signifying that their evidence was inconclusive. The majority of the studies, with exception of Sorenson, Grindstaff, and Turner (1995), utilised a cross-sectional design. In terms of the latter study, the fact that it performed a longitudinal design and utilised some controls denotes that the study can be evaluated under the mediated model. Therefore, the research proceeds to the next stage of the evaluation process to ascertain whether or not there is enough evidence presented to deduce the existence of a relationship. Under the mediated model the publication would receive a rating of A (negative), denoting that the methodology from the publication was conclusive and did not contain any major flaws that would cloud interpretation of the findings; however, the results revealed the negative aspect of church attendance on depression in a sample of teenage mothers.

Good Publications (Rating 7-8)

Within the next subsection 21 (43%) of studies received a high rating of seven to eight, indicating that the design of the studies were above average. The general conclusion produced from these studies is that frequency of attendance at religious services generally provides a protective factor against depression in medical patients. For instance, Al, Dunkle, Peterson, and Bolling, 1998) reported these findings in a sample of 151 medical patients following coronary bypass surgery. The sample was drawn from the cardiac registry at the University of Michigan Medical centre (aged 40-80 years). Psychological adjustment was measured by the Symptoms Checklist-90-R (Derogatis, 1983), which provided a score for general distress. In addition, respondents were asked to respond to a depression subscale for symptoms they experienced during the first month following surgery. Also, information was obtained on health conditions, such as surgical information, post surgical complications, and general health before and after surgery. Perceived social support, though family, friends, and significant others, was measured by the Multidimensional Scale of Perceived Social Support (Zimet, Dahlem, Zimet, and Farley, 1988). Asking patients to respond to questions involving attendance of church services and participating in church activities measured public religious activities. Results from the study revealed that the majority of the sample attended church regularly (54%) and participated in church activities (52%). However, the researchers did not test statistically for the strength of the relationship between the variables of church attendance and depression symptomology. Subsequently, not providing any information for support of the association between these two variables.

However, Koenig, Pargament, and Nielsen (1998) did statistically test for the strength of the relationship between church attendance and depressive symptomology in a sample of 577 medically ill hospitalised medical patients (aged 55 or over). The patients were interviewed from their hospital rooms on questions concerning religious behaviour, such as church attendance, responses ranged from 0 to 3 (0='I've been doing this not at all,' 1= 'I've been doing this a little bit,' 2= 'I've been doing this a medium amount,' 3= 'I've been doing this a lot'). Depression was assessed with an 11-item self-rated depression scale validated in older medical inpatients. Results revealed that church attendance was statistically significant (negatively) with depressive symptoms ($p < .001$), the analysis was controlled for age, race, sex, education, hospital, and medical illness severity.

Similarly, Fernando (1975, 1978) reported findings of a similar nature in a sample of psychiatric patients. The research conducted by Fernando (1975) was a cross-cultural study assessing some familial and social factors in depressive illness amongst psychiatric patients and controls. The study consisted of 117 Jewish and Protestant psychiatric patients and nonpsychiatric controls (from the surgical inpatient service) matched on area of birth and residence (all from the east end of London), social class, age, and sex. The sample was broken down into four groups; 46 Jewish depressives, 71 Protestant depressives, 41 Jewish non-depressives, and 76 Protestant non-depressives. Information on age, social class, marital status, family history of psychiatric illness, birth position and family size, and childhood bereavement was obtained. The Hamilton Rating Scale (Hamilton, 1967) measured the degree of depression. The outcome of the illness was measured by three grades; recovered, improved, and unimproved, through psychiatric notes. Church attendance was utilised

as a measure of religiousness among the Christian samples and religious practices were employed as a measure of religiousness among the Jewish samples. Results from the study revealed that only 4% of depressed Protestants were very religious, measured through church attendance, as compared to 12% of non-depressed Protestants; however, this was not statistically significant ($\chi^2 = 3.18$, $df = 2$, ns). However, depressed Jews were significantly less religious in comparison to non-depressed Jews ($\chi^2 = 9.22$, $p < .01$). The latter study conducted by Fernando (1978) examined the aspects of depression in a Jewish minority group. The study consisted of the same respondents employed in the Fernando (1975) study. Fernando reported from the analysis that 'depression seemed to be associated with a weakening of religious behaviour among Jews-the less orthodox were more likely to become depressed' (p.28). However, this conclusion was not statistically tested for, rather postulated.

The previous studies in this subsection reported on the protective relationship that public religious behaviours had with depression on medical patients and psychiatric patients; similarly, relationships have been reported in differing populations. For instance, Koenig (1995) reported a positive association with older men in prison who reported that they attended Church once per week or more who had significantly lower scores than less frequent attenders (8 vs. 13, $p < .05$). Hertzgaard and Light (1984) reported that women who attended church more than once a month scored lower on the depression scale (in an uncontrolled analysis). Additionally, Brown, Ndubuisi, and Gary (1990) assessed the relationship between religiosity and depression among 451 urban black Americans 18 or over. Church attendance was employed as one of the measures of religiosity as well as three additional items that

measured public religious activities; attending religious crusades, revival meetings, or missions, taking part in religious services, and contributing money to the church. These questions formed part of the Kennedy, Cromwell, and Vaughan (1977) religiosity scale. Mental health was assessed by the Center for Epidemiologic Studies Depression Scale (Radloff, 1977). In addition, the researchers measured stressful life events (Holmes and Rahe, 1967), from this 25-item scale, five most frequently stressful life events were utilised; financial problems, death of a close family member, family members fighting amongst themselves, change in the health of a family member, and personal injury. Socio-demographic data was included in the analysis. Socio-demographic information included age, educational attainment, marital status, employment status, parental status, and residential mobility. Results from the analysis showed that there were significant gender differences in all measures of public religious activities, with females scoring higher ($p < .01$). In relation to depression, although it was observed that a relationship existed with religiosity, as measured by the Kennedy, Cromwell, and Vaughan (1977) religiosity scale, and depression, the measures of public religious activities were not analysed separately prohibiting the inference of the association between this variable and depression.

In a later study assessing 2,956 adults dwelling within the community, Ellison (1995) reported that there were negative correlations between the greater amount of church attendance reported by respondents and lower scores on depressive symptoms, for white respondents ($\beta = -.074, p < .001$), albeit, religious attendance included in the regression analysis alongside the other predictors of religiosity (affiliation and devotion) only explained 1.8% of the variance in depressive symptomology. In spite of this minimal strength, this finding remained consistent, although slightly

weakened, when other covariates were included in the regression analysis ($\beta = -.035$, $p < .01$), with the addition of demographic characteristics, chronic illnesses, stressful life events, and subjective social support. However, for the black respondents, it was reported that there was no association between religious attendance and scores on depressive symptomology.

In addition, few studies reported that there were no association between measures of public religious behaviours, particularly church attendance, and measures of depression. For instance, Brown and Prudo (1981) reported this finding in a sample of 355 women aged 18-65, residing in the rural communities of two islands among the Outer Hebrides off the coast of Scotland. In a later study, Hallstrom and Persson (1984) reported that there was no association between religious attendance and depression in 800 middle-aged women from Sweden. In an elderly sample, Husaini, Blasi, and Miller (1999) reported that there was a positive association between public measures of religiosity and depression, measured by the Center for Epidemiologic Studies Depression Scale (Radloff, 1977), among whites only ($\beta = -.16$, $p < .05$); however, this significant association disappeared whenever medical problems and stress was included in the regression analysis.

The evidence provided by the publications cited by Koenig, McCullough, and Larson (2001) would appear to reveal that there is generally a positive association between public measures of religiosity, in particular church attendance, and lower scores on scales measuring depression, especially in medical and psychiatric samples. However, it generally appears that there is no association between the two variables for women and black respondents. In relation to the A-Levels-of-Evidence strategy, all the above

studies would fall into category C and be eliminated from evaluation as the publications cited utilised a cross-sectional design strategy. Consequently, the evidence presented from the studies are inconclusive ruling out any potential conclusion being inferred surrounding the association between frequent attendance of religious services and depression.

However, within this subsection, few studies reported on the relationship between public measures of religiosity and depression, which utilised a prospective cohort design. For instance, few cross-sectional publications previously stated, revealed the positive associations between church attendance and lower scores of depressive symptomology for medical patients. Similarly, Pressman, Lyons, Larson, and Strain (1990) partially reported on the positive effects of church attendance on depression in 30 elderly women with broken hips (aged 65 and over). The women indicated responses to a variety of questions including depression and the Index of Religiousness (Zuckerman, Kasl, and Ostfeld, 1984), which included an item on frequency of religious attendance. Initial results revealed that when religious attendance was assessed with depression scores, there were no significant association between the two variables; however attendance at religious services was significantly related with discharge depression ($r = -.52, p < .05$). Although, this piece of research incorporated a longitudinal methodology, the researchers did not utilise controls of any nature into their analysis. Therefore, in relation to the systematic review approach the study would be rated as C, denoting that the methodology was not of an adequate standard to receive further consideration.

The next publication potentially of an adequate standard to be evaluated by the systematic review was performed by Williams, Larson, Robert, Buckler, Richard, Heckman, and Pyle (1991). These researchers conducted a study utilising data from Myers, Lindenthal, Pepper, and Ostrander's (1972) and Myers Lindenthal, and Pepper's (1975) longitudinal studies of mental health in New Haven, Connecticut; 720 respondents dwelling within the community, received questionnaires that measured psychological distress, religious commitment, and two summary measures of stressful life experiences. The intention of utilising data from the earlier studies was to assess the extent to which the patterns of findings in the cross-sectional analyses remain robust in the more rigorous prospective analyses. A 20-item symptom checklist scale measuring psychophysiological symptoms indicated the presence of moods of depression and anxiety (Gurin, Veroff, and Feld, 1960) and religious commitment was measured by frequency of church attendance. In addition, socio-demographic factors (age, education, marital status, gender, and race) acted as control variables. Findings report that individuals who state that they attend church regularly report lower levels of psychological distress than infrequent attenders and non-attenders. This relationship remained robust when adjusted for the socio-demographic variables. As in the previous publication (i.e., Pressman, Lyons, Larson, and Strain, 1990), Williams, Larson, Buckler, Heckamn, and Pyle (1991) performed a methodological sound study in respect that they conducted a longitudinal study, employed multiple controls, adequately measured the religion and health variables, and performed statistical analysis. However, they utilised data, which had earlier reports on the same cohort. Consequently, the publication is flawed, under A-Levels-of-Evidence approach, and does not get considered for the next level of the review process.

Excellent Publications (Rating 9-10)

A substantial number of studies (9) were considered of a high standard to be granted a rating of excellent, ranking nine to ten, according to the reviewers. The findings from these studies provide evidence, which is typically unequivocal. The studies from this subsection presented results that represented the positive relationship that public religious activities have on depression, i.e., greater frequency of church attendance is related to lower depressive symptomology. For instance, Ferraro (1998) found that religious practice, as measured by three frequency indicators was negatively correlated with depression, measured by the Center for Epidemiologic Studies Depression Scale (Radloff, 1977), in a sample of 3,497 adults participating in the Americans' Changing Lives 1986 survey ($r = -.01$, $p < .01$).

This positive association between public displays of religiosity and depression was revealed in a sample of elderly individuals. For example, Idler (1987), in a sample of 2,811 elderly persons, revealed that in an uncontrolled analysis, public religiousness, as measured by frequency of church attendance and number of other Congregationalists known to the respondent, was negatively correlated with depression, $R^2 = -.090$, $p < .01$ for males and $R^2 = -.176$, $p < .001$, for females. However, after controlling for multiple controls, such as health behaviours, social support, and coherence (optimism and fatalism), saw the strength of the association being reduced non-significance in men and reduced the association substantially in women ($R^2 = -.079$, $p < .01$).

However, one study within this sub-section revealed no association between public religiousness and depression. In a total sample of 2,679 persons from conservative or mainline Protestant affiliations, Koenig, George, Meador, Blazer, and Dyke (1994) revealed that there was no difference between frequency of church attendance and lower rates of depressive symptomology. The study consisted of 853 'baby boomers' and 1,826 'non-boomers', measured on depression by the Diagnostic Interview Schedule (Robins, Helzer, Croughan, and Ratcliffe, 1981) and a single-item question measuring church attendance. Results showed no differences between high and low church attendance and 6-month or lifetime prevalence rates of depression, controlling for sex, race, health, and socio-economic status.

The typical conclusions that can be postulated from this sub-section are that public religiousness, which generally incorporates a measure of frequency of church attendance, is positively related to depression. It appears that the greater frequency of public religious behaviours is associated with less depressive symptomology, especially within a sample of elderly persons. However, the typical flaw with these studies is that they employ cross-sectional designs. Incorporating a cross-sectional design for a methodology is detrimental for the systematic review process, rendering the evidence presented by the study as inconclusive. Subsequently, the publications are rated as C.

However, few studies within this subsection employed a longitudinal methodology. The findings from these publications are typically consistent. For instance, Idler and Kasl (1992) examined the impact of religion on depression. More specifically, they utilised a single-item question concerning frequency of attendance at religious

services and measured the strength of the relationship between high and low frequency attenders on scores of a measure of depression, Center for Epidemiologic Studies Depression Scale (Radloff, 1977). In addition, numerous confounding variables were measured, such as demographic characteristics, physical health status, health practices, three indicators of social networks, and an indication of an attitude for optimism. At baseline (1982), the sample consisted of 2,812 elderly persons, aged 65 and over, the sample received briefs interviews at one-year intervals, although in 1985 the surviving sample received a full-follow-up interview. The final follow-up was in 1989.

In relation to the variables of interest, church attendance and depression, the results from the analysis revealed that public religiousness had an impact on scores of depression, at the 1985 follow-up when baseline depression was accounted for. However, church attendance was negatively correlated with depression scores ($\beta = -.181$, ns). The variables that were significantly related to depression scores were having a Jewish affiliation ($\beta = -1.679$, $p < .05$), functional disability ($\beta = 1.239$, $p < .001$), change in disability from 1982 and 1985 ($\beta = 1.233$, $p < .001$), and being married ($\beta = 1.859$, $p < .001$).

These findings were comparable to a later empirical study investigating the relationship between public religiousness and depression. Kennedy, Kelman, Thomas and Chen (1996) examined the relation of attendance at religious services to depression, as measured by the Center for Epidemiologic Studies Depression Scale (Radloff, 1977), in a sample of 1,855 older adults. In addition, at baseline, respondents provided information on health, chronic illnesses, physical and cognitive

functioning, utilisation of and attitudes toward health care, interactions with friends, family, and social service agencies, and financial resources. Twenty-four months after baseline, 85% of the original sample completed a second administration of the depression scale and responded to questions regarding changes in problems with activities of daily living, formal and informal support, and whether their health had improved, declined, or stayed the same.

In relation to church attendance and depression scores, results from the analysis revealed that individuals who reported that they did not attend religious services reported higher levels of depression ($\chi^2 = 9.84$, $p < .001$), at baseline. This finding remained consistent after measures of socio-demographic status, immigration, health, and disability, and social support were taken into account. Examining the longitudinal effects of attendance at services and depression, in an uncontrolled analysis, findings revealed that the percentage of persons not attending services was greater in both the depression-emerged group ($\chi^2 = 17.94$, $df = 1$, $p < .001$) and the depression-persisted group ($\chi^2 = 4.33$, $df = 1$, $p < .05$). This effect diminished once the effects of other characteristics were controlled. However, church attendance was not one of the characteristics included in the model distinguishing between respondents who were never depressed from those in whom depression emerged. The characteristics that predicted the emergence of depression included one or more problems with activities of daily living and health becoming worse over the year.

Also Koenig, George, and Peterson (1998) revealed, in a sample of 87 medically ill older patients, that church attendance was non-significantly associated with psychiatric remission. The researchers performed a 47-week longitudinal study on 87

patients admitted to a psychiatric ward, over the age of 60, with a depressive disorder. At baseline, the respondents underwent a 1-2 hour interview providing information on a variety of measures for depression, physical, and psychosocial functioning. The patients also provided information on church attendance, demographics, social support, and treatment factors. After hospital discharge, depressed patients were followed-up at 12-week intervals. The variables that were associated with quicker periods in remission of depression included quality of life, change in functional status during follow-up, family psychiatric history, number of medical diagnoses, social support, and treatment with antidepressants.

Similarly, Musick, Koenig, Hays, and Cohen (1998) validated the former conclusions that public religious behaviours had no effect on depression, in 3-year longitudinal study consisting of 3,007 elderly adults with cancer, aged 65 and over. Respondents indicated information on a variety of variables. Depression was measured by the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) and public religious behaviour was measured by a single-item indicating frequency of attendance. In addition, information on chronic illnesses other than cancer, functional impairment, social interaction, social interaction satisfaction, and socio-demographics were also obtained. Results showed that although service attendance was related inversely to scores on depression, it was not statistically significant.

Regarding all the publications cited within this subsection, it would appear that the findings from the studies reveal a relationship that is significantly related, until the analysis incorporates the control for a number of covariates and protective factors. These variables typically reduce the strength of the relationship to nonsignificance. In

terms of the systematic review the cross-sectional studies would be rated in the category C, as they contain the methodological flaw that prevents inclusion. However, the four studies within this subsection did meet the standards for inclusion into the systematic review. The publications can be rated and the strength of the evidence for the hypothesis, public religious behaviours protect against depression, can be judged.

The first study presented by Idler and Kasl (1992) reported that depression was negatively associated with depression. However, when the analysis was controlled for a number of socio-demographic characteristics and protective factors, the association was reduced to nonsignificance. In relation to the mediated model, the publication receives a rating of A (ns), denoting that the evidence provided by the study was conclusive. The methodology was judged to be sufficiently sound to infer a judgement on the conclusion. However, the study did not support the hypothesis. In relation to the independent model, the publication receives a rating of A (ns), as it doesn't contain any of the flaws that clouds interpretation. However, the study does not lend any support for the hypothesis, which states that public religious behaviours protects against depression. As there is no differences between the mediated model and the independent model it can be suggested that the observed relationship between public religious behaviours and depression can be accounted for the covariates, functional disability, change in disability from 1982 and 1985, and being married.

The second published article presents comparable findings to the latter study. The study performed by Kennedy, Kelman, Thomas, and Chen (1996) reported that at baseline attendance at religious services was significantly related to scores on depression, and this finding remained robust after covariates were controlled. For the

longitudinal associations with depression, it was demonstrated that zero-order statistics reported a positive association between individuals who reported that they never attended religious services and in both the depression-emerged group and the depression-persisted group. However, this effect diminished once the effects of socio-demographic characteristics and health variables were controlled. For the mediated model, the published article gets a rating of A (ns), suggesting that the methodology is conclusive, yet there is no support for the hypothesis. In terms of the independent model, similar findings were reported. The publication receives a rating of A (ns) showing that there was no support for the hypothesis.

Koenig, George, and Peterson (1998) performed the third publication, which can be evaluated under the review process. The study reported that church attendance was not associated with depression in a sample of 87 medically ill older patients. For the models, mediated and independent, the publication obtains a rating of A (ns). Signalling that that the data presents conclusively that there is no support for the hypothesis; public religious behaviour protects against depression.

For the last study within this subsection, it was reported that there was not a significant association between public religious behaviours and depression, in a sample of elderly adults with depression. The study performed by Musick, Koenig, Hays, and Cohen (1998) reported that religious service attendance was not significantly associated with depression, in a controlled analysis. The rating acquired through this publication for the mediated model and the independent model is A (ns). Indicating that the study provides conclusive evidence that there is no support for the hypothesis

The evidence presented by the publications cited by Koenig, McCullough, and Larson (2001) within this subsection would appear to suggest that the relationship between public religious behaviours and depression is one that is typically mixed. It was observed that some of the empirical investigations reported that there were positive associations between public religious measures, particularly religious attendance, whilst other studies reported that there was no association between the two variables. In addition, only one study reported the negative impact that public measures of religiosity had on depression. In relation to A-Levels-of-Evidence approach, it can be seen that the majority of publications cited by Koenig, McCullough, and Larson (2001) were rated as C (inconclusive). This portends that the empirical studies do not meet the minimal methodological standards to be further considered for evaluation. However, one study (Sorenson, Grindstaff, and Turner, 1995), which was presented in the previous subsection, provided a methodology that meet the minimal standards for evaluation; hence, receiving a rating of A (negative) for the mediated model. Similarly, within the latter subsection, there were enough empirical studies that met the minimal standards for evaluation. From these studies, it was revealed that there was no support for the hypothesis, which suggests public religious behaviours protects against depression.

Examining all the subsections within this domain, it was observed that there were enough high-quality publications that were rated, to infer a conclusion based on the data from the studies. Five studies from this section were considered to have met the minimal methodological standards to provide an evaluation of the hypothesis. The five publications were rated in the A category, with four of them providing no support

in the mediated or independent models for the hypothesis. Therefore, the hypothesis that public religious behaviours protect against depression reaches the level of *consistent failures*.

Anxiety

The relationship between religion and anxiety is one that is characterised by much debate by mental health specialists, philosophers, and religious specialists. Some arguing that religion fosters anxiety and neurosis, whilst others contending that religion buffers against and relieves anxiety. Freud (1928) and more recent authors (Ellis, 1980; Rokeach, 1960) have portrayed the religious individual in not so favourable terms, describing them as being inflexible, conforming, prejudiced, and emotionally unstable. These authors speculated that being involved in religion had the tendency to provoke conditions such as anxiety rather than protect against it. The work examining the association between religion and anxiety disorders are minimal. Reviews (e.g., Gartner, Larson, and Allen, 1991) examining the relationship between public religious activities and anxiety, typically reveal that the relationship is ambiguous.

The review under interest, Koenig, McCullough, and Larson's (2001) review, comment that 'The preponderance of evidence suggests that religion as a whole, especially intrinsic religiousness, tends to buffer against anxiety' (p.153). The reviewers believe that there are positive associations between all forms of religiosity and less anxiety, although they do not specifically comment of public religious measures of religiousness. These reviewers point to eighty-one studies that investigate the relationship between some form of religiousness and anxiety.

Weak Publications (Rating 1-4)

From the eighty-one studies, twenty-five of them utilised a measure of public religious activity. A small number of the studies, 15%, were rated in the weak category (1-4) by Koenig, McCullough, and Larson (2001). The findings from two of these three studies revealed that there were no association between the two variables. For instance, Feifel (1974) examined religious conviction and fear of death among the healthy and terminally ill, in a convenience sample of 95 physically healthy and 92 terminally ill patients (both groups similar on age, sex, education, intelligence, marital status, number of children, and recent experiences of death). Frequency of religious attendance was employed as the measure of religiosity and persons were divided into high and low religious groups. The outcome variable was measured by items which assessed frequency of thoughts about death, suicide, and a single-item measuring conscious fear of death. No significant differences were found between those who attended religious services frequently and death anxiety. Similarly, Kutner (1971) investigated the relationship between church attendance and fear of pregnancy, in a sample of 106 women (Mean age 36, white, and middle class). Results from the analysis revealed that church attendance was unrelated to fear of pregnancy.

However, within this subsection, one study reported a positive relationship between public religious activities and anxiety. Martin and Wrightsman (1965) examined death anxiety and church attendance, in a sample of 58 adult members of three churches (Church of Christ, Methodist, and Christian Church). Fear of Death measured the outcome variable. Results from the analysis reveal that church attendance was inversely related to fear of death ($p < .05$).

Although, the majority of the studies within this section report on the lack of an association between public religious activities and anxiety, particular death anxiety, according to A-Levels-of-Evidence strategy, the studies are of inadequate standard to be deemed acceptable to be assessed under the approach. Therefore, the studies go into the inconclusive category and receive a rating of C; hence, they are eliminated from further consideration of the review.

Average Publications (Rating 5-6)

Koenig, McCullough, and Larson (2001), rated the majority of the studies as average in their overall design. The findings from these studies rated in the category, five to six, typically revealed results that were mixed. Few studies in this subsection reported the negative effects of religiosity on anxiety. For instance, Franks, Templer, Cappelletty, and Kauffman (1990-1991) explored death anxiety as a function of religious variables in gay men with and without AIDS. The sample consisted of 51 gay men with AIDS and 64 gay men without AIDS or without HIV+. The respondents filled out a 10-item measure adapted from Templer and Dotson (1970), religious correlates of death anxiety, which included a question on church attendance. The measure of death anxiety was assessed by Templer's Death Anxiety Scale (Templer, 1970). Findings from the analysis revealed that church attendance was positively, albeit weakly, with death anxiety ($p < .05$).

Similar findings, which disclosed the negative effects of religious activities and anxiety were revealed in college students at the University of Alabama (Wilson and Miller, 1968). The study examined fear, anxiety, and religiousness in 100 college

students. Religiousness was measured by self-ratings of church attendance and the outcome variable was anxiety. Findings revealed in an uncontrolled analysis that church attendance was positively correlated, although weakly, with anxiety ($p < .05$).

Few studies, reported the positive effects that public religious activities could potentially have on anxiety. For instance, Aday (1984-1985) examined 181 undergraduate students on a measure of church attendance and Templer's Death Anxiety Scale (Templer, 1970). Results revealed that amongst undergraduates, church attendance was inversely related to death, high church attenders had lower death anxiety than moderate and low attenders ($p < .05$).

In a sample of 60 adults from a community in rural Texas, Spellman, Baskett, and Byrne (1971) reported that regular church attenders were less likely to suffer from anxiety ($p < .001$), measured by the Taylor Manifest Anxiety Scale (Bendig, 1956), compared to respondents who had a history of sudden religious conversions.

Bivens, Neumeyer, Kirchberg, and Moore (1994-1995) examined death concern among 167 gays and bisexual men with varying proximity to AIDS. The sample was split into three groups; 43 respondents who were HIV+, 69 involved as volunteers or professionals in AIDS hospices, wards, or organisations, and 55 gay or bisexual men who were uninvolved with the virus. Death anxiety was the outcome variable and church attendance measured religiousness. Results revealed that church attendance was weakly and inversely associated with less death threat ($r = -.26, p < .05$).

A positive association between public religious activities and anxiety was also reported in research examining death anxiety in terminally ill elderly patients. For instance, Smith, Nehemkis, and Charter (1983-1984) examined a sample of 20 patients on church attendance and four death anxiety scales. Church attendance was inversely related to the first three scales, however, it was only significantly related to only one of the anxiety scales ($r = -.51, p < .05$).

Less anxiety, in relation to frequent church attendance was also reported in a study examining the grief levels of bereaved mothers and fathers (Bohannon, 1991). The study consisted of 143 mothers and 129 fathers who lost a child within the past 18 months. Church attendance measured the public aspect of religiosity and individuals were divided into high and low attenders. Findings revealed for both mothers and fathers that frequent church attendance was inversely related to death anxiety ($p < .01$ and $p < .001$, respectively).

A substantial number of studies within this category reported the lack of an association between public religious activities and anxiety. For instance, Berman (1974) examined the relationship between religion, belief in afterlife, and life threatening experiences. The study consisted of 198 subjects who experienced at least one life-threatening experience (cases) and 198 subjects who reported no near-death experiences (controls). The groups were matched on age, sex, religion, religiosity (devout or moderately devout vs. inactive), and socio-economic status. Results revealed that religiously active subjects were not less likely than inactive subjects to experience anxiety (27% vs. 33%).

This lack of association can also be found in undergraduate students. Hoelter and Epley (1979) studied 375 undergraduate male and female students. Church attendance was employed as one measure of religiosity. Findings showed that church attendance was not related to anxiety. Comparable findings between frequency of church attendance and death anxiety were revealed in a study examining 394 students from four schools, a private university, a public university, a hospital nursing school, and a Catholic seminary college (Minear and Brush, 1980-1981).

Similarly, no association between the two variables were revealed in elderly individuals. Jeffers, Nichols, and Eisdorfer, (1961) asked respondents in a cross-sectional survey of 260 volunteers, aged 60 and over, if they were afraid to die. Results revealed, in an uncontrolled analysis, that religious attendance and frequency of religious activities were not related to fear of death. In a later study examining an elderly sample, Morse and Wisocki (1987), revealed similar results. The researchers utilised a cross-sectional survey in a convenience sample of 156 persons aged 60-90 years. The public religious variable was measured by membership in church/temple and religious attendance, with the majority of the sample being church members (85%). Results revealed significant differences being high and low church attenders ($p < .05$), in an uncontrolled analysis. However, when chronic physical health was controlled, associations with anxiety diminished. Reid, Gilmore, Andrews, and Caird (1978) found no significant associations between groups of church-affiliated members in a sample of 501 persons aged 65 or over. However, this finding was reported due to the lack of statistical analyses being performed on the data.

No association between active religious behaviour and anxiety were also reported in empirical researches examining elderly samples. For instance, Koenig (1988) performed a mail survey on 263 senior ambulatory lunch program participants and 41 home-bound lunch program participants. The questionnaire consisted of items examining religious beliefs and activities, with 1-item assessing active involvement in religious community activities. Death anxiety was assessed through 1-item, with two possible response items, which were combined into a single category. In addition, a number of items assessed subjective general health, stress, and socio-demographic characteristics. Results from a chi-square statistic revealed that there were no significant associations between involvement in religious community activities and anxiety in an elderly sample.

This finding was substantiated in a study examining older adults in a geriatric assessment clinic (Koenig, Moberg, and Kvale, 1988). These researchers surveyed 106 outpatients, aged between 56-94 ($M= 74.4$, $SD= 7.5$). The respondents were asked questions concerning religious beliefs and behaviours, including 2-items on frequency of church attendance and frequency of other church related activities (adult Sunday school, bible study, and prayer group participations). A review of the patient's mental health record revealed the presence or absence of chronic anxiety within a 6-month period. Findings from the study revealed that a high percentage of the sample reported at least weekly church attendance (50%) and 27% of the sample reported weekly participation in other church related activities. In addition, 14.4% of the sample was diagnosed with chronic anxiety within the 6-month period. However, there was no association between the measures on public religious activities and chronic anxiety.

From the section, perceived as average in quality by Koenig, McCullough, and Larson (2001), it would appear that the findings from the empirical research are generally mixed, with the majority pointing to no significant associations between the two variables. The samples that were utilised in the research from this section were diverse, ranging from undergraduate students to terminally ill elderly patients. However, in relation to the strategy placed by the systematic review the evidence from these studies would be deemed as inconclusive and placed in the C category, as they are studies which all incorporate cross-sectional designs, the first elimination approach for this category.

Good Publications (Rating 7-8)

Within the proceeding category cited by Koenig, McCullough, and Larson (2001), two studies were deemed as being above average in their overall design, with the two studies offering conflicting conclusions. The first study performed by Hertsgaard and Light (1984) was on 760 randomly selected females residing on farms in the mideastern state. Results revealed that frequent attenders, measured by those persons who attended church more than once a month and anxiety, measured by Multiple Affect Adjective Checklist (Zuckerman and Lubin, 1965) were significantly associated ($\lambda = .96, p < .001$), after controlling for multiple confounders.

Opposing results to this finding were revealed in a longitudinal study performed by Williams, Larson, Buckler, Heckmann, and Pyle (1991) examining religion and psychological distress in a community sample. Results revealed that church attendance was not associated with anxiety after controlling for psychological distress at baseline. The analysis reported on data from the New Haven mental health study

(Myers, Lindenthal, Pepper, and Ostrander, 1972; Myers, Lindenthal, and Pepper, 1975). At baseline, 720 respondents replied to questions concerning socio-demographic characteristics, as well as, items concerning psychological distress (Gurin, Veroff, and Feld, 1960), and a question asking frequency of church attendance. The Time 2 questionnaire, administered two years later, assessed in addition to the original questionnaire, stressful life events and physical health problems. Findings at Time 2, in relation to church attendance, revealed that church attendance is positively related to psychological distress ($\beta = 0.83$, $p < .001$). This showed that respondents who reported more frequent church attendance score higher on the Gurin scale, revealing that greater church attendance is related to lower scores of anxiety. This finding remained robust after controlling for a number of covariates ($\beta = 0.84$, $p < .001$). However, when baseline scores from the Gurin scale were added into the equation the level of significance reported in the first two models was reduced in non-significance ($\beta = 0.16$, ns). This finding revealed that church attendance was not related to anxiety, as measured by the Gurin scale over a longitudinal period.

It can be postulated that the findings from this subsection showed that church attendance over a period of time is not associated with anxiety. However, anxiety was measured by a scale that measured the collective term psychological distress, which assesses a range of mental health affective disorders, including the presence of moods of depression. Subsequently, the scale is not independently assessing anxiety, therefore stating the null association between church attendance and anxiety is premature. Additionally, in relation to A- Levels-of-Evidence strategy, the studies are eliminated from proceeding to the evaluation stage. The reasoning for this is that the first study utilises a cross-sectional design and the latter study, although it employs a

longitudinal design and uses multiple controls, required its data from an earlier study that utilised the same cohort. Therefore, the evidence provided by the publications within this subsection, have methodological designs that do not meet the minimal standards for evaluation; therefore, receive the rating of C (inconclusive).

Excellent Publications (Rating 9-10).

Within the final category, ranked 9-10, Koenig, McCullough, and Larson (2001) listed two studies that they deemed to be of excellent quality. The studies examined the relationship between religiosity, typically measured through attendance at church services, and anxiety in samples of adults living within the community. The first study performed by Koenig, Ford, George, Blazer, and Meador (1993) utilised a sample from the Wave II of the NIMH Epidemiologic Catchment Area Study. The study consisted of 2,969 adults over the age of 18, stratified into three age groups (young, middle, and old). The measuring instruments administered to the sample included the Diagnostic Interview Schedule (Robins, Helzer, Croughan, and Ratcliffe, 1981), which determined a diverse range of anxiety disorders (6-months and lifetime prevalence rates), using the DSM-III (i.e., simple phobia, social phobia, obsessive-compulsive disorder, panic disorder, agoraphobia, somatization disorder, posttraumatic stress disorder and generalised anxiety disorder).

A substantial amount of additional information was also included such as, socio-demographic and health data. The Duke Social Support Index (Landerman, George, Campbell, and Blazer, 1989) measured four aspects of social support (size of social network, amount of social interaction, instrumental support, and subjective support). Church attendance was measured by a single-item asking respondents to indicate

'About how often do you go to religious meetings or services?' Results from the analyses were stratified by age. First, findings from the younger age (18-39 years) group reveal that for any disorder, there were significant differences in 6-month and lifetime prevalence rates between high and low church attenders (7.5% vs. 12.2%, $p < .01$ 6-month and 12.4% vs. 18.8%, $p < .01$ lifetime, respectively). These differences persisted after controlling for sex, socio-economic status, chronic health problems, and negative life events. In addition, for the young group significant differences existed between low and high church attenders in relation to 6-month prevalence of agoraphobia (5.3% vs. 2.0%, $p < .01$) and lifetime prevalence of general anxiety disorder (13.5% vs. 6.7%). Second, for the middle-aged group (40-59 years), significant differences were found between low and high church attenders, only for 6-month prevalence of social phobia (2.7% vs. 0.2%, $p < .01$). All differences were significant after controlling for sex, socio-economic status, health, and negative life events. No significant differences were revealed for any anxiety disorder for the elderly group (60-97 years). It would appear that religious practices among community-dwelling adults are differentially related to anxiety disorders, with frequent church attendance offering a protective factor for some anxiety disorders, in particular for younger adults.

In the second study, opposing findings were found in relation to elderly adults. Koenig, George, Blazer, Pritchett, and Meador (1993) examined the same cohort of elderly adults as in the previous study, utilising a measure of church attendance. The sample consisted of 1,299 elderly persons aged 60 and over. The same measures employed in the previous research was utilised for the present study. Findings from the correlational analyses between religious variables, socio-demographic, and health

factors revealed that church attendance was significantly and inversely related to age ($r = -.07$, $p < .01$), alcohol abuse and dependence ($r = -.09$, $p < .001$), chronic health problems ($r = -.10$, $p < .001$) and impaired function ($r = -.29$, $p < .001$). In addition, church attendance was found to be positively correlated with sex (if female; $r = .09$, $p < .001$) and social support ($r = .17$, $p < .001$). In relation to the 6-month prevalence of anxiety disorder, church attendance was found to be statistically significant to indicators of anxiety ($\beta = -.11$, $p < .01$). However, religious attendance alongside other constructs of religiousness only accounted for 1% of the total variance. This significant association diminished when mental and health variables were added into the equation. In addition, it was revealed that church attendance was not associated with the lifetime prevalence of anxiety disorders.

The evidence presented by the two publications within these subsections performed cross-sectional designs, therefore, not comprising the proper standards for the review. In order to uncover the complex dynamics of changing religiousness and changing mental health states, longitudinal data is essential. Subsequently, the studies would be rated in the category C (inconclusive). The studies do not get considered for further evaluation.

Initially viewing the published articles cited by Koenig, McCullough, and Larson (2001) it appears that the relationship is typically mixed, with some research showing the positive influence that public religious behaviours has on anxiety, whilst others demonstrate that there is no association between the two variables. However, it must be noted that the majority of the empirical research investigating this branch of religiosity and mental health utilised cross-sectional designs and did not examine the

role of confounding variables or known protective factors. In examining the hypothesis that public religious behaviours protects against anxiety, under A-Levels-of-Evidence approach, it can be viewed that within all subsections, none of the research met the minimal methodological standards to be evaluated. Consequently, the accumulated research reaches the level of *insufficient evidence* for the hypothesis that public religious behaviours are protective against anxiety. This indicates that there is not enough evidence provided for a notable conclusion to be offered. The finding signifies for a greater amount of research of adequate standard to be conducted.

Subjective Well-being

Studies' examining the relationship between church attendance and subjective well-being appears to be a favourable one, with church attendance correlating positively with well-being. It has been reported that church attendance generally has salutary effects on positive affect disorders, including, life satisfaction (e.g., Guy, 1982), happiness (Graney, 1975), and well-being (e.g., Mookherjee, 1994).

Witter, Stock, Okun, and Haring (1985) in a meta-analysis of 28 studies that examined religion (either measured as church attendance or religiosity scores) and subjective well-being, reported that the effect sizes ranged from -.01 to +. 58, with the 95% confidence intervals lying between .14 and .25. In addition, the effect sizes were found to be slightly larger for church attendance (Witter, Stock, Okun, and Haring, 1985). This led to conclusions that "religion is positively associated with perceptions of well-being" (Witter, Stock, Okun, and Haring, 1985, p.355). Therefore, it would

appear that the weight of the evidence suggests church attendance has a small but positive relationship with psychological well-being.

In reality, the findings are frequently mixed and appear to depend largely upon how well-being is measured. For instance, Francis and Kaldor (2002) reported a positive association between church attendance and well-being amongst 989 adults from the Australian community. Well-being was assessed through the Bradburn Balanced Affect Scale (Bradburn, 1969), an instrument that comprises two five-item measures of positive and negative affect. Conversely, Lewis, Joseph and Noble (1996) reported no association between frequency of church attendance and life satisfaction, as measured by the Life Satisfaction Scale (Diener, Emmons, Larsen, and Griffen, 1985) in 150 Northern Irish undergraduate students. However, Markides (1983) found that church attendance was a consistent predictor of life satisfaction, as assessed through the Neugarten's 13-item Life-satisfaction Index (Neugarten, Havighurst, and Tobin, 1961). Church attendance remained significant with sex, age, marital status, health, and years at school partialled out. Therefore, it would appear that the relationship is one, which is typically mixed, depending on the operationalisation of the well-being variable.

In terms of the review offered by Koenig, McCullough, and Larson (2001), the authors pointed to a number of studies that assessed the relationship with public religious activities and well-being, in particular, church attendance. The total number of studies listed by these authors examining this domain was 60, with the authors commenting that a positive relationship between the two variables generally exists.

Weak Publications (Rating 1-4)

In relation to the first ranking category (1-4), the authors listed few studies as being of poor quality (4). This small amount represents 6% of the total within this domain. The typical findings revealed from these studies, is that public religious activities, particularly conceptualised and operationalised as frequent attendance at religious services presents a relationship that is generally mixed with well-being, generally conceptualised and operationalised through measures on well-being, hope and optimism, and self-esteem. For instance, the empirical research presenting no association between the two variables include a study performed by Francis and Bolger (1997) on 55 retired civil servants responding to measures of church attendance and Bradburn Balanced Affect Scale (Bradburn, 1969). A study comparing life satisfaction between 183 aged nuns with 439 other elderly females and 191 elderly males (Kvale, Koenig, and Ferrel, 1989).

One study presented a positive association between the two variables. For instance, McNamara and St. George (1979) sampled 2,164 persons aged 18 or over, and found that church membership was significantly related to personal competence.

The evidence provided by these cited publications would possibly lead to suppositions that the relationship between church attendance and well-being is one that does not exist. However, the major flaw associated with these studies is the cross-sectional methodology employed, which prevents any conclusive statements being deduced. In terms, of A-Levels-of-Evidence approach, the evidence provided by the research would be deemed as inconclusive due to their cross-sectional nature; hence they

would be omitted from the review process. However, within this sub-section assessing the relationship between church attendance and well-being, one of the studies utilised a longitudinal design to investigate the relationship.

The study performed by Farakhan, Lubin, and O'Connor (1984) examined the relationship between life satisfaction and religion among retired black persons, in a 3-month prospective cohort. The study consisted of 30 elderly persons aged from 52 to 97 ($M= 71.4$, $SD= 7.6$) who responded to measures on life satisfaction, and a measure of church attendance. The researchers indicated that church attendance was a primary correlate to life satisfaction, in an uncontrolled analysis (no statistics provided).

Although the above study provided a longitudinal analysis, the fact that the researchers did not incorporate controls in their analysis and did not perform statistical analysis, provides two major flaws that prevents the study from being evaluated by the systematic review. Therefore, the four studies cited within this sub-section are not considered to be of adequate standard by the A-Levels-of-Evidence approach. Accordingly, the studies receive a rating of C (inconclusive).

Average Publications (Rating 5-6)

In relation to the category of studies as seen as of average quality, ranking 5-6, by the reviewers, the general assertion is that public religious activities, in particular, frequent church attendance, is beneficial to well-being. From this section, Koenig, McCullough, and Larson (2001) point to a vast number of publications, 21 in total, which provide results that are generally mixed. The claim, that public religious activities are beneficial to well-being has been substantiated in a number of empirical

studies. For instance, Alexander and Duff (1991) found that public religious behaviours (belonging to a church and frequency of church attendance) was positively related to scores on the Life Satisfaction Index (Liang, 1984) in a sample of 156 retired persons from a religious and secular retirement community, in an uncontrolled analysis ($R^2 = .18, p < .05$). Similarly, Guy (1982) reported greater life satisfaction between the Neugarten's Life Satisfaction Index (Neugarten, Havighurst, and Tobin, 1961) and a measure of church attendance, in a sample of 1,170 older aged persons living within the community. Heisel and Faulkner (1982) further validated this relationship in a sample of 122 urban African Americans, who reported greater happiness and greater life satisfaction scores with greater frequency church attendance. Hunsberger (1995) reported the positive association between a single-item measuring frequency of church attendance and a single-item measure of happiness ($r = .28, p < .01$), in a zero-order correlation for a sample of 85 older adults aged 65 to 88. Morris (1991) reported a moderate correlation between public religious behaviour and well-being in a survey involving 400 elderly persons aged 40 and over, who indicated how active they were in religious activities (church related forms of social participation) and how satisfied they were with life, as measured by a single-item ($r = .19, p < .001$).

The positive association between a single-item measure of church attendance and a single-item measure of general well-being was revealed among 84 caregivers providing care to Alzheimer's disease patients, $r = .34, p < .01$ (Burgener, 1994).

Saunders (1979-1980) reported positive findings in a study comparing 102 adults who were recently bereaved and 107 controls; church attendance was related to greater

optimism. In a later study, which examined the suddenly bereaved, Sherkat and Reed (1992) reported similar findings; church attendance was significantly and positively related to self-esteem, even after controlling for social variables, $R^2 = .19$ ($p < .01$).

Blaine and Cocker (1995) reported positive findings between participation in church services and scores on the Satisfaction with Life Scale (Diener, Emmons, Larsen, and Griffith, 1985), in a sample of 66 African American undergraduate students ($r = .29$, $p < .05$).

This positive association can also be viewed amongst bereaved mothers. Bohannon (1981) surveyed 143 bereaved mothers and 129 bereaved fathers who had experienced the death of a child within the 18 months on a measure of optimism/despair and a measure of church attendance. Results revealed that there were significant differences between low attenders and high attenders on scores of the optimism/despair dimension amongst bereaved mothers but not amongst bereaved fathers.

However, within this subsection few studies did not report the positive associations between public religious behaviour and well-being. Many publications within this subsection reported on the positive relationship between public religious activities, particularly church attendance and well-being in the elderly. However, Steintz (1980) reported that there was no association between the two variables. Similarly no association was revealed in a study examining 156 male students at a Catholic high school (Benson and Spilka, 1973). Results showed that the relationship between hours

spent in church or at church-related activities per month was not significantly associated with scores on a 23-item instrument measuring well-being.

In examining the evidence provided by the publications from this subsection, it could be postulated that there is a positive relationship between public religious activities, particularly church attendance and measures of well-being, particularly in the elderly cohorts. Consequently, the evidence from these studies would not be evaluated under the systematic review process as the researchers performed a cross-sectional methodology, placing the studies in the C category (inconclusive). However, there were few studies within this subsection, which incorporated a longitudinal design. For instance, Carp (1974) performed a prospective cohort study of 133 elders assessed 18 months and 8-years apart to examine adjustment to a new living situation (elderly housing). Happiness was measured by self, peer, and administrators ratings. Results revealed that attendance was not one of the primary predictors of happiness by subjects themselves.

This finding was substantiated in a later study assessing the effects of age on well-being. Graney (1975) assessed the relationship between happiness and social participation in the elderly. This longitudinal study consisted of 60 elderly women aged 62 to 89 years, at baseline. After four years the remaining forty-six subjects who survived were re-interviewed. The subjects were split into three groups according to age range, the youngest group consisted of 16 persons aged 66 to 75, the middle group consisted of 13 persons aged 77-81, and the oldest group consisted of 17 persons aged 82-92 years. Respondents were interviewed and asked to self-report their level of happiness by indicating on a 10-item scale feelings of positive and negative affect;

revised from the Bradburn Affect Balance Scale (Bradburn, 1969). A question on frequent attendance of religious services was utilised as a measure of social participation. No controls were employed in the study. Results from the analysis reveal that at baseline frequent attendance at religious services was positively associated with happiness ($p < .01$), as measured by the revised scale. Additionally, this effect remained at the four-year follow-up, that is frequent religious attendance was related to happiness ($p < .01$). When the group was stratified for age (young, middle, and old), this differential relationship remained, both at baseline and after the four-year follow-up (all at the $p < .01$ significance level).

Although, the studies reported on associations between church attendance and well-being, by utilising a longitudinal design, the researchers did not employ controls within their analyses. Therefore, it is not possible to infer if the relationship is entirely affected by the religious variables. It could be that the religious variable is a proxy measure for some second alternative variable. For instance, church attendance could be acting as a proxy measure for social support. Failure to control for these factors can lead to biased estimations of the association. Subsequently, the latter studies receive the rating of C (inconclusive).

Good Publications (Rating 7-8)

Within this subsection, in which the publications were considered above average in their overall design, ranking 7-8. The reviewers cited twenty-three empirical publications. Similarly, in opposition to the previous subsections, the relationship between the two variables is one, which is generally positive. The majority of the empirical research describes a positive relationship between the public religious

activities and well-being. For instance, Edwards (1973) reported that church-related activities were positively associated with scores on a 10-item life satisfaction scale in a sample of 507 predominantly white persons over the age of 45 ($\beta = .14, p < .05$).

Similarly, Spreitzer and Synder (1974) reported in a sample of 1,447 respondents that church attendance was weakly related to scores on a single-item of life satisfaction ($r = .08, p < .01$). However, when the respondents were stratified by age it appeared that church attendance was related to life satisfaction for individuals under the age of 65 ($r = .12, p < .05$), but not for respondents over the age of 65. Additionally, when socio-economic status was controlled for, the significance value was reduced in non-significance.

In a later study surveying 6,621 adults, Gee and Veevers (1990) reported that religious activity was related to scores on a 7-item life satisfaction scale, involving six life domains for both men and women aged from 25 to 59, controlled for sex. However, in a cross-sectional survey that controlled for multiple confounders, Lee and Ishii-Kuntz (1987) found in a sample of 2,872 respondents over the age of 55 that church attendance was positively related to scores on a 7-item measure of morale, $r = .12$ ($p < .01$) for males and $r = .17$ ($p < .01$) for females. However, when a number of covariates were entered into the analysis, the significant association remained only for men. Usui, Keil, and Durig (1985) also revealed that church attendance was positively related to scores on the Neugarten's Life Satisfaction Index (Neugarten, Havighurst, and Tobin, 1961) in a sample of 704 persons aged 60 and over, after controlling for a number of socio-demographic covariates ($\beta = .26, p < .01$).

In a later study this finding was substantiated; Koenig, Kvale, and Ferrel (1988) reported in a sample of 836 persons over the age of 55, that public religious activities (church attendance and other religious group-related activities) were positively related to subjective well-being, as measured by the Philadelphia Geriatric Center Morale Scale (Lawton, 1972). A zero-order correlation showed that the variables were moderately related ($r = .26, p < .01$). When the effects of known covariates were controlled in the sample (sex, age, and health), the association remained statistically significant, albeit weakened ($\beta = .18, p < .001$).

Comparable findings have been disclosed in subsequent studies assessing the elderly. For instance, Coke (1992) reported that church participation (hours per week) were related to scores on the Satisfaction with Life Scale (Diener, Emmons, Larsen, and Griffin, 1985) for males ($r = .38, p < .01$) but not females in a sample of 166 elderly African Americans. However, this was a zero-order correlation. However, McGloshen and O'Bryant (1988) revealed in sample of widowed females that greater frequency of church attendance was positively and significantly related to scores on Bradburn's Balance Affect Scale (Bradburn, 1969), $r = .19, p < .01$.

Equally, beyond a 'normal' population, it has been reported that public religious activities is positively related to well-being. For instance, in a sample of 1,174 drug addicts, receiving treatment, Hater, Singh, and Simpson (1985) reported, in an uncontrolled analysis, that church attendance was significantly associated with well-being ($r = .14, p < .01$).

Research has also been conducted that explored the relationship between church attendance and well-being, between blacks and whites. The study performed by St. George and McNamara (1984) examined a cohort of 3,877 white persons and 449 black persons on a range of items measuring life satisfaction and single-item measure of frequency of church attendance. Results revealed that for the male white respondents, church attendance was weakly related to global happiness ($\beta = -.14$, $p < .01$) after controlling for age, education, income, and occupational prestige.

Similar results were revealed by St. George and McNamara (1984) for the female white respondents, after controlling for the covariates. Church attendance was related to scores on the global happiness measure ($\beta = -.12$, $p < .01$). However, a stronger relationship was observed between church attendance and global happiness for the black male and female respondents, $\beta = -.25$ ($p < .001$) and $\beta = -.27$ ($p < .001$), respectively. This finding reveals that for both races and genders, there is a relationship between greater frequency at church services and higher scores on a global happiness measure. However, looking at the R^2 statistics for the groups it was shown that church attendance only explained about 2% of the extra variance in global happiness scores over and above the covariates for the white sample, but for the black sample it explained on average 6% of the extra variance, over and above the covariates.

There were few studies within this subsection, which did not report on any significant association between the two variables. It has been postulated that there are differences between individuals from different races, on how church attendance relates to well-being. However, Ortega, Crutchfield, and Rushing (1983) examined race differences

in personal well-being, in a sample of 4,522 persons from the community. The respondents were asked to provide information on a variety of socio-demographic variables, such as race, education, family income, age, marital status, sex, health status and place of residence. In addition, respondents replied to measures on frequency of social interaction (including church attendance) and three measures of personal well-being (happiness, general life satisfaction, and relative life satisfaction), which were measured by responses to a single-item. Findings from a controlled analysis revealed the greater well-being, as measured by the three items of well-being, was not predicted by church attendance.

In a few studies, it was revealed that public religiousness was related to certain dimensions of well-being, whilst not being significantly correlated with other dimensions of well-being within the same cohort. For instance, Pollner (1989) examined the relationship between divine relation, social relations, and well-being in a sample of 3,072 adults involved in the 1983 and 1984 General Social Survey. Respondents indicated information on frequency of church attendance, marital status, race, sex, family income, age, education, recent traumatic events (life stressors), and social interaction. Respondents also indicated their level of satisfaction by responding to questions on general well-being, which included dimensions on life satisfaction, global happiness, marital happiness, and life excitement. Results from a zero-order regression revealed that church attendance was significantly related to global happiness ($R^2 = .005$, $p < .01$) and marital happiness ($R^2 = .014$, $p < .01$), but not significantly related to life satisfaction or life excitement. The analysis between church attendance and the measures of well-being did not incorporate controls, the measure of church attendance was investigated as a background variable.

Similarly, Poloma and Pendleton (1990) reported that church attendance was related to some dimensions of general well-being, whilst not being related to others. The study performed by the researchers examined the relationships between religious domains, including church attendance and general well-being (life satisfaction, negative affect, existential well-being, and happiness), in a sample of 560 respondents interviewed by telephone. In a zero-order correlation results revealed that church attendance was significantly related to life satisfaction ($r = .21, p < .01$), happiness ($r = .12, p < .05$), and existential well-being ($r = .23, p < .01$), but not to negative affect ($r = -.06, ns$).

Within this subsection, few of the researchers reported that the relationship between public religiousness and well-being in the elderly remained significant, albeit weakened, when covariates were added in the analysis (i.e., Koenig, Kvale, and Ferrel, 1988; Lee and Ishii-Kuntz, 1987). Levin and Markides (1988) provide partial support for this finding in a sample of 375 elderly and 375 middle-aged respondents from Mexican American cultures. Levin and Markides (1988) performed the study to examine whether religious attendance had a significant effect on psychological well-being net of confounding variables (age, marital status, social class, and health status). The respondents were asked to indicate their frequency of religious attendance and respond to questions on life satisfaction, measured by the 13-item version of the Neugarten's Life Satisfaction Index (Neugarten, Havighurst, and Tobin, 1961). In addition, respondents provided information on age, marital status, education, income, and a standard measure of subjective health and functional health. Results from a zero-order regression revealed that scores on religious attendance and scores on the

well-being scale were positively correlated for older men ($\beta = .21, p < .05$), older women ($\beta = .18, p < .01$), and middle-aged women ($\beta = .23, p < .001$), but not for middle-aged men ($\beta = .06, ns$). However, this was a zero-order analysis, not controlling for covariates. A partial regression, holding for age, marital status, income, education, and subjective health, weakened the association for the older women ($\beta = .13, p < .05$) and middle-aged women ($\beta = .17, p < .01$) and reduced the association to nonsignificance for older aged men. The same findings were revealed in a second analysis this time controlling for the background variables and functional health; for the women cohorts the strength of the association was reduced, but remained significant for the male cohort it was reduced to nonsignificance.

Examining the results from the studies, it would appear that a potential relationship exists between public religiousness, particularly religious attendance and well-being. The inclusion of demographics and particularly health variables weaken the association, particularly in women and tend to reduce the significant association to nonsignificance in men. The vast numbers of studies that have been cited within this subsection have come from a cross-sectional methodology, hence reducing the ability to deduce any suppositions of cause and consequence between the variables. However, within the subsection, Koenig, McCullough, and Larson (2001) cited three published studies that performed a longitudinal methodology.

From the three longitudinal studies, it was revealed that the relationship was one that was equivocal; the results were typically mixed. For instance, Blazer and Palmore (1976) investigated the correlates of religion and happiness, usefulness, and adjustment in a sample of 272 community residents aged 60-94. The sample

contained both white and black persons, male and female participants that approximated the race and sex distribution of the population from which the sample was drawn (Blazer and Palmore, 1976). The Activity dimension of the Chicago Inventory of Activities and Attitudes (Burgess, Cavan, and Havighurst, 1948) included a measure of church attendance, along with several additional measures, which captured private religious activities. The Attitude dimension of the scale measured the variable 'happiness'. In addition, data on sex, race, age, and socio-economic status was obtained. Results from the study revealed that religious activities generally were associated with happiness and usefulness. Religious activities were significantly related to happiness ($r = .16$), especially among men ($r = .26$) and persons over the age of 70 at baseline ($r = .25$). However, none of these correlations were controlled for. In relation to the longitudinal effects of religious activities in relation to happiness, it was stated that the correlations tended to remain robust especially among elderly persons (no statistics provided).

Although, the former study performed a longitudinal methodology, the analysis of the data was typically zero-order; therefore, it is not possible to infer the operative role that religion plays on happiness. In addition, the authors measured church attendance in conjunction with private measures of religiosity; not establishing which variable is responsible for the correlation between religious activities and depression. In terms of A-Levels-of-Evidence approach, the study would be omitted from the systematic review, as it failed to control for known additional variables that may affect the relationship, receiving the rating of C (inconclusive). However, the subsequent studies employ longitudinal methodologies and incorporate controls into their analyses.

The first study conducted by Markides (1983) examined whether church attendance was related positively to positive adjustment (life satisfaction) net of other important predictors of adjustment, in a 4-year longitudinal study of 338 (510 persons at baseline) Mexican-Americans and Anglos aged 60 and over. Church attendance was based on a 6-point item, ranging from every week to never. Positive adjustment was measured by the Neugarten's 13-item version of the Life Satisfaction Index (Neugarten, Havighurst, and Tobin, 1961). In addition, information on age, sex, race, marital status, years at school, health, and socio-economic status was indicated. Results from a regression analysis revealed that at baseline, church attendance for Mexican-Americans and Anglos were significant predictors of life satisfaction despite the inclusion of other important variables ($\beta = .14$, $p < .05$ and $\beta = .19$, $p < .05$, respectively). At Time 2, this finding remained consistent; church attendance was positively related to life satisfaction ($\beta = .12$, $p < .05$ and $\beta = .38$, $p < .05$).

The second study performed by Markides, Levin, and Ray (1987) employed the same cohort as the former study. The sample was measured at three time points over an 8-year period using the same measures as Markides (1983). At baseline, the sample contained 511 Mexican-American and Anglo participants, at Time 2, this number was reduced to 338 participants, and at Time 3, this amount was further reduced to 254 participants. An analysis of the 254 participants across three time bands revealed that church attendance was not related to life satisfaction scores in a sample of Mexican-Americans and Anglo participants. However, when the all available respondents were employed at each time point there was a significant relationship between church attendance and life satisfaction, i.e., $\beta = .09$ ($p < .05$) for the 511 respondents available at baseline and at Time 2, $\beta = .15$, $p < .01$ for the 338 respondents. This finding could

denote the importance of the potential effect of dropouts on an analysis of longitudinal study.

Although, the studies utilised longitudinal designs, incorporated controls, provided statistical analyses, and adequately measured the religion and well-being variable; only one of the studies could be utilised for the systematic review process, as both studies use the same cohort. Consequently, the study that utilised the longest period of time is included in the review. The rating that the second study would receive by utilising A-Levels-of-Evidence approach would be an A, meaning that the methodological of the study conducted by Markides, Levin, and Ray (1987) was conclusive, and no important methodological limitations were identified, which would cloud the interpretation. In this case, the publication received a rating of A (ns). Indicating that the study did not provide support for the hypothesis.

Excellent Publications (Rating 9-10)

In the final subsection of this category, Koenig, McCullough, and Larson (2001) list a plethora of empirical studies, which are considered excellent quality. Therefore the studies were ranked in the top category, 9-10. The typically findings from these studies is that public religious activities presents a mixed relationship with well-being, with some of the research stating that there is a positive relationship between the variables. For instance, Ellison, Gay, and Glass (1989) reported that the frequency of attendance at religious meetings was significantly related to general life satisfaction, after a number of covariates were included in the model ($\beta = .08, p < .05$), for a sample of participants involved in the 1983 NORC General Social Survey. In a survey examining the determinants of satisfaction in both 597 black and 5032 white persons;

Thomas and Holmes (1992) found that higher scores on attendance at religious services and higher scores on a religiously-minded item (summed to provide a composite score) were positively related to scores on a 11-item scale of satisfaction of life, for whites ($r = .17, p < .001$) and for blacks ($r = .29, p < .001$). After 13 covariates were controlled for in the analysis, the associations remained robust between the two variables, although, the association was somewhat more strongly related to blacks ($\beta = .17, p < .001$) than for whites ($\beta = .06, p < .001$).

Similarly, Ellison and Gay (1990) reported, in a controlled analysis, that frequent attendance at religious services was significantly related to well-being, as measured by a single-item 4-level global life satisfaction item ($\beta = .065, p < .05$), in a sample of 2,107 African-American adults. Comparable findings were revealed in a publication, which examined the relationship of how religious behaviour effects the life satisfaction among black Americans. The study performed by Levin, Chatters, and Taylor (1995) utilised data from the National survey of black Americans, a nationally representative cross-sectional survey of the black adult population aged 18 or over. Responses from 1,848 subjects were used in the study and information on age, gender, employment status, marital status, residence (rural or urban), years of education, and income were obtained. A 5-item measuring battery assessed the latent construct, organisational religiosity. The items included; 'How often do you attend religious services', 'Are you an official member of a church or other place of worship', 'How many church clubs or organisations do you belong to or participate in', 'Besides regular service, how often do you take part in other activities at your place of worship', and 'Do you hold any positions or offices in your church or place of worship'. The latent construct, Life satisfaction, was measured by 3-items, 'In general

how satisfied are you of your life as a whole these days', 'taking all things together, how would you say things are these days', and 'Up to now, have you gotten mostly what you hoped for out of life or have you gotten less than you hoped for'.

Results from the analysis, using the LISREL package, revealed that the data was a satisfactory fit for the proposed model. Subsequently, it was shown that organisational religiosity, as measured by the five constructs, were positively related to the 3-item measure of life satisfaction ($\beta = .28, p < .01$). This significant effect remained after controlling for the seven demographic constructs, such as age, gender, education, marital status, employment status, geographical region (Southern vs. non-Southern), and area of residence (urban vs. rural); $\beta = .26 (p < .05)$.

In addition, positive relationships between public religious activities and subjective well-being, were expressed in the elderly. For instance, Singh and Williams (1982) performed a cross-sectional survey of pooled data from six national probability samples (1973-1978) on 1,459 persons aged 65 and over. It was reported that those persons who indicated that they attended religious services more frequently were much more satisfied with their physical health than those who attended less frequently ($p < .001$). Although, eleven additional covariates were included in the data, they were not controlled for in the analysis between religion and health.

However, few studies revealed that public religious activities were not related to life satisfaction. For instance, Ellison (1991) reported that frequency at religious services was not related to two indices of subjective well-being (life satisfaction and personal happiness) from data taken from the 1988 General Health Survey. Data was utilised

on a variety of socio-demographics, church attendance and subjective well-being. Results revealed that there was no association between the two variables when socio-demographics were included into the regression equation. The strongest predictors of life satisfaction were education ($\beta = .849$), existential certainty ($\beta = .681$), income ($\beta = .120$) marital status ($\beta = .111$), and belonging to a nondenominational Protestant group ($\beta = .128$). The best predictors of personal happiness were existential certainty ($\beta = .410$), marital status ($\beta = .191$), Catholic ($\beta = -.120$), and traumatic life events ($\beta = -.120$).

This finding was also revealed in a study performed by Krause (1992), who examined psychological well-being in relation to religious attendance in a sample of 448 older African-Americans over the age of 60. Religious attendance was measured by a single-item indicator that reflected how often study participants attended religious services. The construct, subjective well-being was measured by Rosenberg's Self-Esteem Scale (Rosenberg, 1965). In addition, information was obtained on physical health, demographics (age, sex, and education), and social support. Findings revealed that attendance at religious services was not related to self-esteem.

Similarly, Krause (1993) reported, in a sample of 709 persons over the age of 55, that the two variables were not statistically associated. Public religious activities were measured via 3-items; 'Do you belong to a church or religious organisation', 'Do you currently do any unpaid voluntary work for a church or religious organisation', and 'Apart from weddings, funerals, and baptisms, about how often do you attend religious services.' Similarly, a 3-item measuring battery asked respondents to rate

how satisfied they were with their lives. Results showed that the items measuring public religious activities were not related to life satisfaction.

The findings from these publications within this subsection would typically reveal that public religious activities, such as attendance at religious services has salutary associations with life satisfaction in samples of adults dwelling within the community and of elderly persons dwelling in the community. Similarly, findings were also established that pointed to a lack of an association between the two variables from parallel samples. However, the methodology incorporated within these publications are from cross-sectional designs; therefore, it is not possible to postulate whether religiosity is a cause or consequence of a given level of psychological functioning. Due to this limitation, the rating that the publications received within this subsection is C (inconclusive). Favourably, one publication employed a longitudinal methodology and in addition, met the criteria that were set for inclusion for evaluation by the systematic review process.

The study conducted by Levin, Markides, and Ray (1996) examined the relationship between religious attendance and psychological well-being in Mexican Americans. The longitudinal study was performed over an 11-year period, in which three generations of Mexican Americans were investigated. Information obtained for the study included; a single item measure of religious attendance, three summary scales measuring dimensions of psychological well-being (life satisfaction, depressed affect, and positive affect), six exogenous constructs (socio-demographic variables and a single item measure of self-rating health status). Life satisfaction was measured by the Life Satisfaction Index (Neugarten, Havighurst, and Tobin, 1961) and the four-item

subscale of the Center for Epidemiologic Studies Depression Scale (Radloff, 1977) measured positive affect.

Findings revealed that at Time 1, religious attendance was significantly associated with life satisfaction for the oldest and middle-age generations ($\beta = .23$, $p < .01$ and $\beta = .17$, $p < .01$, respectively). Although, this strength weakened at Time 1 when the exogenous construct, were controlled, $\beta = .19$ ($p < .05$) and $\beta = .16$ ($p < .01$). In addition, there were no longitudinal effects of Time 1 religious attendance on Time 2 life satisfaction. It was demonstrated, by examining the scale measuring positive affect, that the longitudinal effects of religious attendance had a positive association with the variable for the youngest generation only ($\beta = .11$, $p < .05$). However, when the control for the exogenous variables were included in the analysis, this significant association remained consistent, albeit negative ($\beta = -.13$, $p < .05$). The researchers suggested that the negative net effect of one of the exogenous variables (education) on Time 2 positive affect contributed to this anomalous finding. However, a second explanation is that positive affect in general, or that particular measure, is simply not reflective of the many more valid well-being measures in existence.

It would appear that the findings from this study provides partially support for the hypothesis that public religious behaviours increase subjective well-being; however, due to the irregularities of the finding with positive affect, it is not possible to infer a definite conclusion. In examining the study under A-Levels-of-Evidence approach, it can be proposed that the study is considered to meet the methodological standards to be evaluated under the review process. The study performed a longitudinal design, controlled for potential confounders, adequately measured the religion and health

variables, and performed statistical analyses. First, the study will be evaluated under the mediated model; the publication did not contain any flaws, thus receives a rating of A, denoting the methodological provides conclusive evidence. However, the data presented nonsignificant results, signifying that there was no support for the hypothesis.

Investigating the published articles cited by Koenig, McCullough, and Larson (2001), it seems that the relationship is one that is typically mixed. Religious attendance has been found to be associated cross-sectionally with measures of happiness (Heisel and Faulkner, 1982) and morale (Lee and Ishi-Kuntz, 1987), but nonsignificant findings have also been reported (Ellison, 1991). However, utilising an objective strategy to evaluate the studies it is possible to extract which of the studies meet the minimally acceptable methodological standards that could depict the direction of the relationship. From the plethora of studies cited in this section, only two meet the standards stated by the systematic review to be evaluated. In relation to evaluating the strength of the hypothesis; public religious behaviours increases subjective well-being. In examining the ratings provided by the evaluation, there appears to be enough high-quality studies to infer a conclusion. The two publications were rated in the A (ns) category. Thereby, suggesting that there was conclusive evidence to suggest that there was no association between the two variables. Subsequently, the hypothesis resulted in repeated failures to support it, the evidence is described as *consistent failures*.

Conclusion

The review conducted by Koenig, McCullough, and Larson (2001) postulates the association between public measures of religiosity (religious attendance) and mental health (anxiety, depression, and subjective-being) in a favourable direction. It has been well-documented by these reviewers that public measures of religiosity, in particular frequency at religious services, are associated with the protection of mental ill health and with the elevation of greater feelings of subjective well-being.

Earlier reviewers (e.g., McCullough and Larson, 1999) remark that the data on public religious behaviours and depression supports the conclusion that depression is indeed less prevalent among people who are involved in public religious activities. However, reviewers (e.g., Gartner, Larson, and Lyons, 1991) state that the data examining public religious behaviours and anxiety is less consistent, with the empirical research typically presenting conflicting findings. In terms of subjective well-being, Levin (1989) concluded that after reviewing two dozen studies linking religiosity to well-being, instruments measuring public religious behaviours exert positive effects of well-being. It appears that the majority of the authors investigating these domains agree that on the most part the relationship, public displays of religiosity, boasts salutary effects on constructs measuring mental health. However, in closer examination of the evidence it has been demonstrated that the relationship is mixed and even contradictory findings tend to be the rule.

In examining depression, anxiety, and subjective well-being it appears that the relationships are typically unrelated. It has been found that zero-order correlations mark the existence of a relationship between public religious behaviours and each of

the mental health variables, but when exogenous constructs are added to the analysis this finding is diminished (e.g., Koenig, Moberg, and Kvale, 1988; Kroll and Sheenan, 1989).

This chapter reassessed the publications that were cited by Koenig, McCullough, and Larson (2001), and found that the claims that the reviewers were remarking on in their review was generally overstated. In subjecting the studies to a objective and systematic approach, it was observed that the publications were fraught with many methodological flaws that needed addressing before any conclusions could be validated.

It was observed that there was enough evidence available to address the general conclusion that public religious behaviour, such as religious church attendance, offers protection from depression. Examining this domain, it was revealed that there were five high-quality publications, which were rated in the A category. Four of the studies (Idler and Kasl, 1992; Kennedy, Kelman, Thomas, and Chen, 1996; Koenig, George, and Peterson, 1998; Musick, Koenig, Hays, and Cohen, 1998) provided no support for the mediated or the independent models for the hypothesis. The findings from the four studies reached the level of *consistent failures*.

In relation to the cited publications, which examined the relationship between public religious measures and anxiety, it can be suggested that the relationship was one that was generally mixed. When the publications are subjected to the rigorous methods of the systematic review, it was clearly observed that none of the publications met the minimum standards to be evaluated to ascertain whether or there was enough

evidence to support that hypothesis. Therefore, the studies that were cited by Koenig, McCullough, and Larson (2001) reached the level of *insufficient evidence*.

It was demonstrated that there were substantive amounts of research conducted into the area of public religious behaviours and subjective well-being. Examining the publications within this domain revealed that the preponderance of findings showed that there was no association between the two variables, especially when covariates were controlled in the model. However, when the publications were scrutinised under the systematic review process, it was ascertained that the majority of the research was methodologically inadequate to be evaluated. Within, the subsections, two studies were revealed (Levin, Markides, and Ray, 1996; Markides, Levin, and Ray, 1987) that did meet the standards. From these publications it was demonstrated that evidence for the hypothesis reached the level of *consistent failures*.

The three mental health issues that have been substantially examined by empirical researchers do not meet the minimal methodological standards to be evaluated by an objective and systematic review process. From the sections that did contain enough high-quality research articles, the conclusion, according to A-Levels-of-Evidence strategy, was that there was no support for the hypothesis, which states that public religious activities benefit mental health. It could be suggested that there is too much optimism expressed by researchers and authors interested in this area of study or maybe there is not enough high-quality empirical research conducted on the subject to communicate a definite conclusion.