

Summer 2022

Social Network Analysis of Online Support Communities for Adolescent and Young Adult Cancer Survivors

Carlos Perez Aldana

University of Texas Health Science Center at Houston, Carlos.A.PerezAldana@uth.tmc.edu

Follow this and additional works at: https://digitalcommons.library.tmc.edu/uthshis_dissertations



Part of the [Biomedical Informatics Commons](#)

Recommended Citation

Aldana, Carlos Perez, "Social Network Analysis of Online Support Communities for Adolescent and Young Adult Cancer Survivors" (2022). *Dissertations (Open Access)*. 56.

https://digitalcommons.library.tmc.edu/uthshis_dissertations/56

This is brought to you for free and open access by the McWilliams School of Biomedical Informatics at DigitalCommons@TMC. It has been accepted for inclusion in Dissertations (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digcommons@library.tmc.edu.

Social Network Analysis of Online Support Communities for Adolescent and Young Adult Cancer Survivors

by

Carlos Arturo Pérez Aldana, M.S.

APPROVED:

DocuSigned by:

Tiffany Champagne-Langabeer

Tiffany Langabeer, Ph.D., Chair

DocuSigned by:

Susan H. Fenton

Susan H. Fenton, Ph.D.

DocuSigned by:

Sahiti Myneni

Sahiti Myneni, Ph.D.

Date approved: 7/25/2022

Social Network Analysis of Online Support Communities for Adolescent and Young
Adult Cancer Survivors

A
Dissertation

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
School of Biomedical Informatics
in Partial Fulfilment of the Requirements for the Degree of
Doctor of Philosophy

By

Carlos Arturo Pérez Aldana, M.S.

University of Texas Health Science Center at Houston

2022

Dissertation Committee:

Tiffany Langabeer, Ph.D.¹, Advisor
Susan H. Fenton, Ph.D.¹
Sahiti Myneni, Ph.D.¹

¹The School of Biomedical Informatics

Copyright by
Carlos Arturo Pérez Aldana
2022

Dedication

I dedicate this work to the courageous, young people that were faced with their cancer battles prematurely.

I also dedicate this effort to my strongest support system, my wife and son, Kristi and Thomas, whose love and encouragement drove me throughout my doctoral journey...Thomas, I met my deadline!

Acknowledgements

I am especially grateful to my committee. Dr. Langabeer, my advisor, has provided me with the motivation, energy, and knowledge throughout this arduous endeavor. Her unconditional support has been key to complete this work. Dr. Fenton has been a role model to follow, she showed me goals can be reached with dedication, hard work, and discipline. Dr. Myneni has provided me with insightful and methodological guidance. Her trust in my sometimes-meticulous ways allowed me to approach this work with the methods that made my efforts enjoyable. My classmates provided me with great companionship and continuous support to keep going throughout my journey.

My parents, to whom I will be forever indebted, provided me with their unconditional love, support, teaching and opportunities. They contributed extensively to me becoming who I am and getting to where I am.

Abstract

There are an estimated 633,000 adolescent and young adult (AYA) cancer survivors in the U.S. and nearly 89,500 AYAs are diagnosed with cancer every year. Cancer creates developmental and life stage disruptions, which result in multiple survivorship challenges, particularly among AYAs. Despite the advances made in cancer oncology and survivorship care, AYA cancer survivors continue to face diverse and unique psychosocial needs. Research suggests that online support communities have the potential to positively impact psychosocial care by providing AYA cancer survivors with access to social support which can help them successfully transition from treatment back to normal life as well as improve their well-being. In addition, online support communities have become important sources of social support, particularly peer support, offering an opportunity for AYA cancer survivors to exchange support and overcome psychosocial challenges. However, despite an increasing use of online support communities by cancer survivors in general, there is limited evidence providing insights into how online social support can be leveraged by AYA cancer survivors to bridge existing gaps in their psychosocial care. This study provides a deeper understanding of online support exchange by examining the structures of support networks of online interactions among AYA cancer survivors. It applies an informatics approach that combines content analysis, computerized text analysis, and social network analysis. The

results show that AYA cancer survivors are mostly exchanging emotional support but also exchange informational and esteem support in similar proportions. In addition, this study expands current understanding of how AYA cancer survivors are using language to exchange support online. Furthermore, the structural characteristics of support networks reveal they are characterized by low densities and average degrees. Moreover, subcommunities of network support developed among AYA cancer survivors, in spite of low levels of cohesion and clustering between them. Additionally, support networks show that AYA cancer survivors who exchange informational or esteem support are also likely to exchange emotional support. Lastly, the novel data-driven insights gathered by applying an informatics approach may inform the future design and implementation of online support interventions that aim to address the unmet psychosocial needs of AYA cancer survivors.

Vita

2008Bachelor's Degree in Electronic
Engineering, El Bosque University, Bogotá, Colombia

2008M.S. (Specialization), Project Management,
El Bosque University, Bogotá, Colombia

2010M.S., Telemedicine and eHealth,
Polytechnic University of Madrid, Spain

2016 to present.....Graduate Research Assistant, Teaching
Assistant, Ph.D. Student, School of Biomedical Informatics, The University of Texas
Health Science Center at Houston, Houston, TX

Publications

- Langabeer, J. R., Henry, T. D., **Perez Aldana, C.**, DeLuna, L., Silva, N., & Champagne-Langabeer, T. (2018). Effects of a Community Population Health Initiative on Blood Pressure Control in Latinos. *Journal of the American Heart Association*, 7(21), e010282. <https://doi.org/10.1161/JAHA.118.010282>
- Pérez-Aldana, C. A.**, Lewinski, A. A., Johnson, C. M., Vorderstrasse, A. A., & Myneni, S. (2021a). Exchanges in a Virtual Environment for Diabetes Self-Management Education and Support: Social Network Analysis. *JMIR Diabetes*, 6(1), e21611. <https://doi.org/10.2196/21611>
- Pérez-Aldana, C. A.**, Lewinski, A. A., Johnson, C. M., Vorderstrasse, A. A., & Myneni, S. (2021b). Unpacking Peer Conversations in a Virtual Community for Diabetes Self-Management Education and Support: Behavior Science and Linguistics Perspective. *Studies in Health Technology and Informatics*, 284, 25–30. <https://doi.org/10.3233/SHTI210655>
- Singh, T., **Perez, C. A.**, Roberts, K., Cobb, N., Franklin, A., & Myneni, S. (2019). Characterization of Behavioral Transitions Through Social Media Analysis: A Mixed-Methods Approach. *Studies in Health Technology and Informatics*, 264, 1228–1232. <https://doi.org/10.3233/SHTI190422>

Field of Study

Biomedical Informatics

Table of Contents

Abstract	iv
Vita.....	vi
Publications.....	vii
Field of Study.....	viii
Table of Contents	ix
List of Tables	xii
List of Figures	xiii
Chapter 1: Introduction	1
Chapter 2: Literature Review	3
2.1 AYA Cancer Survivors Psychosocial Needs	5
2.2 AYA Cancer Survivors and Social Support.....	9
2.3. Online Support Communities Targeting AYA Cancer Survivors	12
2.4 Online Social Support and Social Network Analysis	16
Chapter 3: Materials.....	20
3.1 Research Online Support Community	20
3.2 Data Collection and Sample.....	23
Chapter 4. Specific Aim 1: To Characterize the Content of Online Interactions Among AYA Cancer Survivors in Terms of Types of Social Support, Survivorship Care Plan Domains, Psychosocial Need Themes, and Linguistic Features.....	26

4.1 Objective 1: Develop a Coding Scheme to Determine the Presence and Extent of Types of Social Support, Survivorship Care Plan Domains, and Psychosocial Needs in Online Interactions among AYA Cancer Survivors	27
4.1.1 Methods.....	28
4.1.2 Results.....	40
4.1.3 Conclusion	43
4.2 Objective 2: Determine the Presence and Extent of Types of Social Support, Survivorship Care Plan Domains, and Psychosocial Need Themes of Online Interactions Among AYA Cancer Survivors	43
4.2.1 Methods.....	44
4.2.2 Results.....	44
4.2.3 Conclusions.....	53
4.3 Objective 3: Extract the Linguistic Features of Supportive Online Interactions Among AYA Cancer Survivors	55
4.3.1 Methods.....	56
4.3.2 Results.....	60
4.3.3 Conclusions.....	66
Chapter 5. Specific Aim 2: To Examine Network Structures of Online Interactions Among AYA Cancer Survivors Based on Content Characteristics Determined in Specific Aim 1	68
5.1 Objective 1: Characterize Social Network Structures Based on Types of Social Support, SCP Domains, and Psychosocial Need Themes.....	69
5.1.1 Methods.....	69
5.1.2 Results.....	73
5.1.3 Conclusions.....	81

5.2 Objective 2: Assess the Relationships Between the Structure of Networks Based on Types of Social Support and Linguistic Features	82
5.2.1 Methods.....	82
5.2.2 Results.....	85
5.2.3 Conclusions.....	88
Chapter 6: Discussion and Conclusions.....	89
6.1 Discussion for Specific Aim 1: Characterization of the Content of Online Interactions Among AYA Cancer Survivors in Terms of Types of Social Support, Survivorship Care Plan Domains, Psychosocial Needs, and Linguistic Features	89
6.2 Discussion for Specific Aim 2: Examination of Network Structures of Online Interactions Among AYA Cancer Survivors Based on Content Characteristics Determined in Specific Aim 1	93
6.3 Limitations	95
6.4 Conclusions.....	95
6.4.1 Innovation	96
6.4.2 Contributions to Biomedical Informatics.....	96
References.....	98
Appendix A.....	108
Appendix B.....	109
Appendix C	112
Appendix D.....	125
Appendix E	129

List of Tables

Table 1. Coding scheme overview	40
Table 2. Descriptive statistics for messages containing social support categories.....	44
Table 3. Descriptive statistics for messages containing SCP domains categories	47
Table 4. Descriptive statistics for message author categories	53
Table 5. Network structure measures	71
Table 6. Summary of network metrics for CSN Young Adult forum's social support networks	78
Table 7. QAP correlation results	85

List of Figures

Figure 1. Screen capture of the CSN site	21
Figure 2. Frequency distribution of number of messages per topic	24
Figure 3. Frequency distribution of number of messages per participant	25
Figure 4. Content analysis process	28
Figure 5. Number of messages by types of social support requested and provided	46
Figure 6. Proportions of messages that requested and provided social support aligned with SCP domain categories	48
Figure 7. Number of messages that contained social support aligned with SCP domain categories	49
Figure 8. Frequency of psychosocial need themes	50
Figure 9. Frequency of psychosocial need themes by types of social support	51
Figure 10. Frequency of psychosocial need themes within SCP domains	52
Figure 11. Mean scores for LIWC summary measures for AYA cancer survivors' online exchanges that requested and provided social support	61
Figure 12. Mean percentage of total words for LIWC psychological categories of affective and social processes by exchanges that requested/provided social support.....	63
Figure 13. Mean percentage of total words for LIWC psychological categories of cognitive processes and drives by exchanges that requested/provided social support....	64
Figure 14. Communication support network.....	74
Figure 15. Informational support network	76
Figure 16. Emotional support network.....	76
Figure 17. Esteem support network	77
Figure 18. Tangible aid (left) and network (right) support networks	77

Figure 19. Two-mode network for the SCP domain 1 cancer surveillance and screening, formed by CSN discussion board “Young Cancer Survivors” participants and psychosocial need themes in their support exchanges	79
Figure 20. Two-mode network for the SCP domain 2 late effects/side effects management, formed by CSN discussion board “Young Cancer Survivors” participants and psychosocial need themes in their support exchanges	80
Figure 21. Two-mode network for the SCP domain 3 risk reduction and cancer prevention, formed by CSN discussion board “Young Cancer Survivors” participants and psychosocial need themes in their support exchanges	80
Figure 22. Two-mode network for the SCP domain 4 psychosocial functioning, formed by CSN discussion board “Young Cancer Survivors” participants and psychosocial need themes in their support exchanges	81
Figure 23. Correlation matrix for support providing exchanges showing the correlations between linguistic features and network structural properties	87
Figure 24. Correlation matrix for support requesting exchanges showing the correlations between linguistic features and network structural properties.....	88

Chapter 1: Introduction

This study applies an informatics approach to investigate online social support exchanges among adolescent and young adult (AYA) cancer survivors. I hypothesize that a deeper understanding of the structures of online support networks and the linguistic features of the content of such interactions among AYA cancer survivors will provide insights into online support communities aimed at this subpopulation. These insights may inform the design of impactful online support interventions that meet AYA cancer survivors' psychosocial needs. This study addresses the following twofold research question:

1. What are the content characteristics of online interactions among AYA cancer survivors in terms of: (a) types of social support, (b) survivorship care plan (SCP) domains, (c) psychosocial needs, and (d) linguistic features?
2. Based on the identified content characteristics, what are the properties of the structure of online support networks, and how are these structural network properties associated with: (a) types of social support, (b) SCP-aligned psychosocial support needs, and (c) linguistic features?

To answer these research questions, this study addresses the following specific aims:

Specific Aim (SA) 1. To characterize the content of online interactions among AYA cancer survivors in terms of types of social support, survivorship care plan domains, psychosocial needs, and linguistic features. SA 1 entailed the following objectives:

- a. Develop a content analysis methodology based on social support theory, cancer survivorship care plan models, and psychosocial needs of cancer survivors.
- b. Determine the presence and extent of types of social support and SCP-aligned psychosocial support needs by manually coding online interactions among AYA cancer survivors.
- c. Extract the linguistic features of online interactions among AYA cancer survivors.

SA 1 applies content analysis and computerized text analysis. The content analysis is informed by a theoretical framework that combines social support theory, cancer survivorship care plan models, and psychosocial needs of cancer survivors.

Specific Aim 2. To examine network structures and properties of online interactions among AYA cancer survivors based on content characteristics determined in SA 1. This entailed the following objectives:

- a. Reveal and characterize social network structures based on types of social support and SCP-aligned psychosocial support needs.
- b. Assess the relationships between the structure of networks based on types of social support and linguistic features.

SA 2 applied a social network analysis approach in order to reveal, characterize, and assess the structure of online support networks via quantitative network measures and statistical methods.

Chapter 2: Literature Review

Cancer is one of the main public health problems the world faces today. In the United States, cancer remains the second leading cause of death with a projected 609,360 deaths expected to occur in 2022 (Siegel et al., 2022). In addition, while the overall number of cancer diagnoses in men has decreased between the period of 1990-2013 and stabilized through 2018, and whereas in women it has been stable through 2010 with a slight increase in recent years, more than 1.7 million Americans were diagnosed with cancer in 2018 alone (Centers for Disease Control and Prevention [CDC], 2021; Siegel et al., 2022). Recent cancer epidemiological data indicate that more Americans are surviving cancer. The 5-year relative survival rate for all cancers combined has increased from 49% to 68% between the mid-1970s and 2011 through 2017 (Siegel et al., 2022). Consequently, the number of cancer survivors in the U.S. reached an estimated 16.9 million in 2019 and projections indicate this number will grow to more than 22.1 million by 2030 (Miller et al., 2019). Yet, despite these epidemiological trends indicating the progress made in early detection and treatment of cancer, the burden associated with cancer is substantial and set to spiral upward (Park & Look, 2019; Yabroff et al., 2021). Research is increasingly documenting the economic burden of cancer exposing the impacts on survivors and their families, their caregivers, and society as a whole. Cancer survivors have substantial medical and non-medical costs, including higher out-of-pocket and time costs, as well as productivity costs compared to those without cancer (Lorgelly & Neri, 2018; Park & Look, 2019; Yabroff et al., 2021). For instance, health care expenditures for cancer survivors are nearly four times higher than those without cancer

(Park & Look, 2019). Furthermore, Essue and colleagues (2020) showed that 44% of cancer survivors are impacted by psychosocial costs, which can lead survivors and their families and caregivers to poorer health, clinical, and economic outcomes (Essue et al., 2020). Lastly, cancer care expenditures by the U.S. healthcare system were projected to reach to \$173 billion by 2020 and are expected to continue increasing in the near future (Park & Look, 2019).

The present study focuses on adolescent and young adult (AYA) cancer survivors. While a cancer survivor was defined more than three decades ago as “a patient who has had cancer [...] from the time of diagnosis through the remainder of his or her life” (Rodriguez & Lewis-Patterson, 2019, p. 3), the Adolescent and Young Adult Oncology Progress Review Group (AYAO PRG, 2006) more recently defined the AYA cancer survivor population with upper and lower age limits as the population “comprising individuals aged 15 through 39 years at cancer diagnosis” (AYAO PRG, 2006, p. 3). Thus, hereafter, an AYA cancer survivor will refer to an individual who was diagnosed with cancer between ages 15–39 through the remainder of his or her life. Current estimates indicate there are more than 633,000 AYA cancer survivors in the United States and approximately 89,500 AYAs were diagnosed with cancer in 2020 alone (Chao et al., 2020; Miller et al., 2020). Moreover, the overall 5-year relative survival rate among AYAs has equally increased since the mid-1970s to an estimated 85.5% (National Cancer Institute [NCI], n.d.). Finally, an increasing body of literature in AYA oncology and survivorship indicates AYA cancer survivors require specific urgent attention. AYA cancer survivors have been recognized as a patient subgroup characterized by unique

medical and psychosocial needs (Barnett et al., 2016; Close et al., 2019; Jin et al., 2021; Miller et al., 2020). The psychosocial needs AYA cancer survivors face throughout their cancer trajectories are described next.

2.1 AYA Cancer Survivors Psychosocial Needs

AYA cancer survivors are a special population with unique medical and psychosocial needs. AYAs experience unique and complex life phases including adolescence, emerging, and young adulthood. Such life phases are characterized by diverse physical, emotional, cognitive, and social transitions (Janssen et al., 2021). When AYAs are diagnosed with cancer, these typical and already challenging life stages are abruptly disrupted, significantly impacting the achievement of key age-specific developmental milestones (Janssen et al., 2021; Perez et al., 2020). Typical AYA developmental milestones include, but are not limited to, “establishing autonomy, moving away to college, developing romantic relationships, launching a career, becoming financially independent, and starting a family” (Perez et al., 2020). Thus, AYA cancer survivors’ developmental milestones set their needs apart from those of younger and older survivor populations (i.e., pediatric and older adult survivors) as well as their cancer-free peers (Janssen et al., 2021). As a result, AYA cancer survivors are faced with unique and diverse psychosocial needs during and after cancer treatment (Janssen et al., 2021; Jin et al., 2021; Institute of Medicine [IOM], 2008).

The IOM recognized in 2008 that cancer effects are influenced by the physical and developmental age of patients and caregivers while emphasizing that physical and psychosocial stressors created or exacerbated by cancer are intertwined, both resulting

from and affecting each other (IOM, 2008). Consequently, the needs AYA cancer survivors face are complex and diverse due to their multifactorial nature, commonly including developmental, physical, psychological, and social factors, among others (Essue et al., 2020; IOM, 2008). The literature has increasingly documented during the last decade the needs of AYA cancer survivors. Studies show AYA cancer survivors face a myriad of psychosocial needs that can be grouped into four domains: physical, psychological, spiritual, and social (Essue et al., 2020; IOM, 2008). On the one hand, cancer and its treatment effects often lead to disability and overall poor physical health. AYA cancer survivors are faced with challenges in regard to chronic conditions, neurocognitive deficits, fertility, sexual dysfunction, altered body image, and physical condition (Janssen et al., 2021; Jin et al., 2021). For instance, many AYA cancer survivors experience long-term side effects related to cancer treatment, including increased risk for secondary malignancies, cardiovascular disease, endocrine dysfunction, memory and attention deficits, fatigue, and pain (Janssen et al., 2021; Jin et al., 2021; IOM, 2008). In addition, AYA cancer survivors receive little information on fertility prognosis and preservation by the medical professionals despite being an important consideration among this population (Jin et al., 2021). Furthermore, AYA cancer survivors are extremely vulnerable to the stigma of an altered body image, which substantially affects their psychological and social well-being (Jin et al., 2021).

On the other hand, undergoing a cancer diagnosis, its treatment, and survivorship creates substantial levels of distress and exacerbates preexisting psychosocial stressors among AYA cancer survivors. In addition, physical and psychological impairments

caused by cancer often lead AYA cancer survivors to significant social problems. As a result, AYA cancer survivors can be faced with a variety of psychological, spiritual, and social issues (IOM, 2008). First, psychological issues include psychological distress (involving depression and anxiety), post-traumatic stress, and fear of cancer recurrence (Janssen et al., 2021; Jin et al., 2021; Perez et al., 2020). AYA cancer survivors have a higher risk of developing depression and anxiety, and fear of recurrence is significantly greater compared to older survivors (Janssen et al., 2021). Moreover, the COVID-19 pandemic worsened the feelings of anxiety, isolation, and uncertainty among AYA cancer survivors, highlighting their need for psychosocial support (Shay et al., 2021). Second, major social issues include challenges related to education, employment, financial independence, and relationships. AYA cancer survivors undergo education and work disruptions due to cancer treatment. While returning to school or work is key for survivorship, work is also essential to preserving health insurance. Moreover, returning to school or work is associated with quality of life and returning to “normal life” (Jin et al., 2021). Additionally, the financial impact of cancer is substantial and can even be more stressful than experiencing a cancer diagnosis. Whereas young adults are trying to find health insurance coverage other than their parent’s plans, cancer may prevent them from independent health insurance, a situation that can lead them to financial toxicity. Hence, young adults experience depression, stress, and anxiety due to financial toxicity (Janssen et al., 2021; Jin et al., 2021; Perez et al., 2020). Furthermore, AYA cancer survivors often report impaired relationships with friends and family as well as romantic relationships. Maintaining connections and communicating with friends and family members can be

challenging given that AYA cancer survivors feel they are treated differently and misunderstood, emphasizing the complexities of communicating as a cancer survivor. In addition, dating and romantic relationships are a source of worry and uncertainty for AYA cancer survivors (Choi et al., 2022).

Evidence shows the diversity and complexity of the psychosocial needs AYA cancer survivors are faced with, yet these needs remain largely unmet among many AYA cancer survivors (Janssen et al., 2021; Jin et al., 2021; Perez et al., 2020; Stepan & Waldmann, 2019). Despite the progress made in recent years in addressing the psychosocial needs among cancer survivors overall, AYA cancer survivors continue to report a lack of age-appropriate resources and services for most of their psychosocial needs, particularly information and social support (Jin et al., 2021; Lea et al., 2020; Perez et al., 2020). Lastly, emerging research suggest that AYA cancer survivors face many barriers when accessing psychosocial care. Holland and colleagues (2021) identified that AYA cancer survivors face person-centered, service-related, and systemic factors-related barriers when accessing psychosocial care (Holland et al., 2021). Studies show guidelines for the management of long-term treatment effects, monitoring for secondary malignancies, preventative care, or patient-provider communication targeting AYA cancer survivors are lacking (Hydeman et al., 2019). In conclusion, AYA cancer survivors continue to face a myriad of psychosocial issues which remain largely unaddressed and require urgent attention (Hydeman et al., 2019; Jin et al., 2021; Lea et al., 2020; Perez et al., 2020).

2.2 AYA Cancer Survivors and Social Support

Meeting the psychosocial needs of AYA cancer survivors is key given the significant impact they pose during cancer treatment and survivorship. The psychosocial issues that cancer creates or exacerbates often lead to multifactorial consequences. Accordingly, the IOM (2008) stressed the negative role those psychosocial stressors play in the morbidity, mortality, and functional status of cancer survivors. Psychosocial stressors commonly include mental health problems such as distress and depression, limited access to resources such as financial resources, and inadequate social support (IOM, 2008). In addition, psychosocial stressors often lead to suffering, diminished quality of life, and poor health outcomes. Studies show AYA cancer survivors often feel lonely and experience social isolation throughout their cancer journey (Jin, 2021; Choi, 2022). Isolation among AYA cancer survivors can result from the difficulty keeping up with peers due to the age-specific disruptions created by cancer, personal feelings that act as barriers when seeking support from friends or family, and the lack of perceived and actual support received (Jin, 2021). Furthermore, the distress created by psychosocial stressors can negatively affect the cardiovascular, immune, and endocrine systems. Psychosocial issues not only affect cancer survivors, but also their families and the community overall (IOM, 2008). However, there are some psychological traits that have been shown to protect against psychosocial distress, causing positive health effects among cancer survivors. Particularly, social support is one of the psychosocial factors with the strongest evidence for positive benefits on health (Cohen & McKay, 1984; Goldsmith, 2004; IOM, 2008; Thoits, 2011; Wright, 2016).

Several theoretical frameworks have been developed and used to study social support in both face-to-face and online contexts (Wright, 2016). The key theoretical frameworks commonly used in the study of social support include the buffering effect model (Cobb, 1976), the optimal matching model (Cutrona & Russell, 1990), the social comparison theory (Festinger, 1954), the social information processing theory (Walther, 1992), and the strength of weak ties theory (Granovetter, 1973; Wright, 2016). The optimal matching model, Cutrona & Russell's theoretical framework (1990), has been used extensively by researchers, particularly in the study of online social support (Cutrona & Russell, 1990; Wright, 2016). Their social support typology (1990) defines five types of support: informational, emotional, esteem, network, and tangible (Cutrona & Russell, 1990). Moreover, research shows these types of social support are commonly exchanged between members of online support communities (OSCs) for diverse health purposes (Egbert & Wright, 2019). OSCs have been characterized as “topic-driven platforms where people with diverse backgrounds and perspectives can address individual needs or specific questions as a collective” (Abendschein, 2020). OSCs offer advantages such as anonymity, easy access, and availability which has positioned them as a key source of social support to manage anxiety and uncertainty, or address knowledge gaps, among other psychosocial issues (Abendschein, 2020; Egbert & Wright, 2019).

Social support is a common coping strategy used by cancer survivors to deal with their illness. In addition, studies show that family and friends can be the main support sources among AYA cancer survivors (Penn & Kuperberg, 2018). Furthermore, support groups have been shown to create a sense of safety that may increase the willingness of

AYA cancer survivors to share their stories to foster belonging and normalcy. Peer support groups, including online support groups, can help survivors cope with their illness (Jin, 2021). Peer support has been long used as a therapeutic strategy, and cancer survivors have received benefits from it for a few decades (Zebrack & Isaacson, 2012). Studies show peer support provides support in unique ways and can be a more effective coping strategy. In addition, studies show peer support can help AYA cancer survivors with developmental and psychological adjustment as well as to reduce distress and anxiety. Moreover, meeting peer survivors and peer support meeting peer survivors was identified as one the most important supportive needs among AYA cancer survivors (Jin, 2021; Choi, 2022; Penn, 2018).

AYA cancer survivors are also using readily available social media and online support communities to connect with peers and find support (Jin, 2021). Thus, OSGs are also an important source of support. Early studies using content analysis of OSCs showed emotional and informational support were more often exchanged among members of OSCs than companionship and tangible support (Braithwaite et al., 1999; Hwang et al., 2010, 2011). Furthermore, evidence from content analyses of OSCs suggests that action-facilitating types of support (i.e., informational and tangible support) are more common among OSCs aimed at patients with chronic conditions such as cancer (Rains et al., 2015). Based on existing OSC research studies, social support has the potential to impact psychosocial care positively so that AYA cancer survivors successfully transition from treatment to survivorship (Barnett et al., 2016; Docherty et al., 2015; Galán et al., 2018; Masterton & Tariman, 2016; Nass et al., 2015). However, despite the increasing use of

OSCs by cancer survivors, there is very limited evidence on how OSCs can be leveraged to bridge the existing psychosocial care gaps in AYA cancer survivorship (Chou & Moskowitz, 2016; Galán et al., 2018; Masterton & Tariman, 2016; Nass et al., 2015). Online social support is thus key to helping AYA cancer survivors receive optimal survivorship care so that they can achieve their full potential (Barnett et al., 2016; Docherty et al., 2015; Galán et al., 2018; Masterton & Tariman, 2016; Nass et al., 2015; Warner et al., 2016).

2.3. Online Support Communities Targeting AYA Cancer Survivors

Extensive research exists on OSCs aimed at cancer survivors. A review of the literature shows that most research on OSCs for cancer survivors has focused on two major areas: characterization of these communities (including their members and their behaviors) and impact of participation (Zhang et al., 2017). In addition, evidence suggests that participating in OSCs benefits cancer survivors in different ways. Researchers have shown that by participating in OSCs, cancer survivors increase their access to social support (Rodgers & Chen, 2005), reduce levels of depression and stress (Beaudoin & Tao, 2008), and are better equipped to cope with their cancer journey (Maloney-Krichmar & Preece, 2005).

The increasing engagement of individuals with diverse health needs in OSCs is encouraging. OSC use by people facing health problems has increased during the last two decades providing a large amount of data available to researchers (Wright, 2016). These interactions are rapidly increasing the availability of user-generated content, which has enormous potential to advance the understanding of online social support. Consequently,

informatics techniques and tools have become crucial to study OSCs and how online social support is exchanged among members of these communities. Moreover, novel informatics approaches have been used to advance online cancer survivor support research. Initial research focusing on message content consisted of topic and sentiment analyses of cancer forums using computational approaches based on machine learning and text mining techniques (Qiu et al., 2011; Portier et al., 2013; Zhao et al., 2014). Early attempts to identify social support embedded within online interactions from online cancer survivor communities included similar computational approaches. Wang and colleagues (2012) used machine learning models to automatically classify informational and emotional support embedded within online interactions from an online breast cancer forum (Wang et al., 2012). In addition, Biyani and colleagues (2014) explored the use of binary machine learning-based text classifiers to identify emotional and informational support in an OSC for breast cancer survivors at scale (Biyani et al., 2014). Similarly, based on Biyani and colleagues' work, Wang, Yen, & Reitter (2015) used the same binary machine learning-based classifier for their study on linguistic alignment; they identified emotional and informational support at scale in an OSC for survivors of breast and colorectal cancer (Wang, Yen, & Reitter, 2015). A different approach to identifying emotional and informational support at scale was used by Wang, Kraut, & Levine (2015). They used MTurks (Amazon Mechanical Turk) to manually code a sample of interactions and then applied machine learning techniques to model the support dynamics of eliciting and providing emotional and informational support in online support communities (Wang, Kraut, & Levine, 2015). Lastly, another approach requiring the identification of

social support at scale from an Online Health Community (OHC) for cancer survivors was used by Wang, Zhao, & Street (2017). They applied text mining techniques to study associations between types of support (emotional and informational) and user participation (Wang, Zhao, & Street, 2017). In summary, a review of the literature shows limited online support research focusing on message content. In addition, the few studies that have explored the identification of support types in OSCs have been limited to only emotional and informational types of support (Wang et al., 2012; Biyani et al., 2014; Wang et al., 2015; Wang et al., 2015; Wang et al., 2015; Wang et al., 2017). Despite evidence suggesting that emotional and informational types of support are more frequently exchanged among members of OSCs in general, research on OSCs for cancer survivor support specifically has shown other types of social support (esteem/appraisal, network, instrumental/tangible support) are also being exchanged amongst its members (Bambina, 2007).

Considering that online social support was not the main research focus for all studies described above, a gap in knowledge and a current opportunity exists to study all types of support being exchanged between members of OSCs. Additionally, despite the potential for informatics to advance research on OSCs and online social support, studies focusing specifically on AYA cancer survivors using informatics approaches are recent and sparse. The literature shows there are only a few studies on OSCs for AYA cancer survivors using modern informatics methods and tools. Computerized text analysis has recently been used to study the language used by AYA cancer survivors to communicate and interact with other survivors within OSCs. Crook and colleagues (2016) revealed

significant linguistic differences between online support soliciting posts that were replied to (N=342) and those that were not (N=342) in an OSC for YA cancer survivors. Linguistic differences included emotions, cognitive processes, and pronoun use, among others. In addition, online posts that were replied to were characterized by being briefer and having more first-person pronouns, negative emotions, and present and past tense verbs (Crook et al., 2016). Another study by Thompson and colleagues (2016) compared the language used by AYA cancer survivors when communicating either online or face-to-face to discuss diverse topics such as health, work, and leisure, among others (Thompson et al., 2016). They found significant differences in the communications' content and style words (Thompson et al., 2016). In addition, AYAs communicating via an online discussion board used more future tense and more words relating to friends, sex, anger, sadness, causation, and inhibition. Thompson and her colleagues' findings suggest that OSCs facilitate communications regarding friendships, sex, negative emotions, and future-oriented questions and concerns among AYA cancer survivors (Thompson et al., 2016). More recently, Warner and colleagues (2018) analyzed Instagram posts discussing young adult (YA) cancer themes with the objective of describing social support, sentiment, and linguistic features embedded in communications from this highly popular social media (Warner et al., 2018). Results showed differences within YA cancer-related posts, specifically, by treatment status (active treatment or survivorship) and type of user (individual or organization). YA cancer survivors received more social support (using the number of "likes" as a proxy indicator of support), used more positive and emotional terms, and fewer pronouns than users in active treatment.

Also, individuals gained more social support (again, using the number of comments as a proxy indicator of support) and used more positive terms, but less emotional terms than organizations (Warner et al., 2018). Finally, Kaal and colleagues (2018) analyzed a sample of 1,896 posts from an OSC for Dutch AYA cancer survivors (Kaal et al., 2018). In addition to emotional support being the most exchanged type of support among OSC members, Kaal and her colleagues found linguistic features showing that community members used a higher number of words indicating cognitive, social, and affective processes. Interestingly, community members showed insight gains as revealed by the words they used to communicate (Kaal et al., 2018).

2.4 Online Social Support and Social Network Analysis

Social Network Analysis (SNA) is an approach increasingly used in many disciplines, including biomedical informatics. In contrast to the content of online supportive interactions, a structural or social network perspective focuses on how social ties or interactions provide social support and thus predict outcomes (Hether, Murphy, & Valente, 2016). A SNA perspective thus explains associations between network characteristics and social support (Hether, Murphy, & Valente, 2016). Researchers have suggested that combining the support content with the social networks can provide in-depth social support analyses (Hether, Murphy, & Valente, 2016). A thorough review of the literature identified only one study focusing on an online support forum for cancer survivors which combined the support content of interactions with social networks (Bambina, 2007). Bambina (2007) identified types of social support by manually coding the content of 1,149 online interactions (posts) according to informational, emotional and

companionship support. Companionship was the type of support more frequently exchanged among members of the forum, followed by emotional and informational support (Bambina, 2007). Additionally, social network analyses revealed an unbalanced relationship between member roles of support “givers” or “takers”, where givers exchanged all types of support while takers received informational support only and did not reciprocate with any support type. Furthermore, network-level analyses revealed key network location or actors: stars, prime givers, serious members, and moderate users. Such key actors were associated with various levels of support exchange according to network location and level of cohesiveness. Going further, in addition to analyzing the relationship between network associations and social support exchanges, a combined approach allowed Bambina (2007) to assess how and why social support was disseminated throughout the social network.

The literature review also identified three health-related studies that combined the support content of interactions with social networks but focused on non-cancer topics: psychosis online support, pregnancy and prenatal health online support, and online smoking cessation (Chang, 2009; Hether, Murphy, and Valente, 2016; Zhang and Yang, 2015). These studies are relevant because of the combination of methods used by the researchers. These methodological approaches informed the present study in different ways. First, Chang’s (2009) study combined content analysis with a network approach to analyze supportive interactions from an online psychosis support community (Chang, 2009). Content analysis involved the manual coding of 558 posts to identify types of support according to Cutrona and Russell (1990) and Cutrona and Suhr (1992) social

support typology (informational, esteem, network, and emotional) and an additional “thanks” category due to “a large amount of posts containing a gesture of gratitude” (Chang, 2009). Results from Wang’s (2009) analysis showed network and informational support were exchanged most frequently amongst members of the online community. In addition, network characteristics allowed an in-depth analysis of communication patterns of five uniplex and one multiplex support networks (Chang, 2009). Second, Hether et al. (2016) manually coded 525 messages seeking support and 1,965 messages providing support from two OHCs for pregnancy and prenatal health according to Cutrona and Suhr (1992). They found interactions requesting informational and emotional support were most frequent. In addition, a functional approach combined with a SNA approach allowed them to extend previous research by showing that support network structures varied across support types (Hether et al., 2016). Third, Zhang and Yang (2015) analyzed support exchange patterns and behaviors in the QuitStop forum (Zhang and Yang, 2015). In addition, they explored support exchanges by smoking quit stages. Results showed that informational support was exchanged most frequently among members of the community. Furthermore, Zhang and Yang (2015) found that “givers of informational support have been abstinent for a longer time than givers of nurturant support, and receivers of informational support have been abstinent for a shorter time than receivers of nurturant support” (Zhang and Yang, 2015). Finally, further related research has highlighted the benefits of combining SNA and the content to better understand online social support exchange and OSCs in general (Moessnet et al., 2018; Myneni et al., 2013, 2015, 2016). In summary, a review of the literature shows SNA can provide in-depth

analyses of the social dynamics caused by the exchange of online social support.

However, only one study (Bambina, 2007) has combined both support content and social network approaches to the study of online support communities for cancer survivors, and no studies have examined online support exchanges among AYA cancer survivors.

Chapter 3: Materials

3.1 Research Online Support Community

The Cancer Survivors Network (CSN; <https://csn.cancer.org/>), launched in July 2000 by the American Cancer Society (ACS), is a non-commercial and publicly available online support community (OSC) that provides cancer survivors and caregivers with a private and secure way to find and connect with peers. CSN's purpose is to facilitate the exchange of experience-based knowledge and peer support and all the site's content is contributed by its members (CSN, n.d.). CSN offers its members discussion boards, individual member pages where they can share their stories, private messaging, and a chat room. As of January 2022, CSN has 28 community discussion boards or forums including cancer type specific and other discussion boards. Cancer specific forums are focused on types of cancer (e.g., leukemia) while other discussion boards focus on diverse topics such as humor and caregiver concerns, to name a few. Both community discussion board categories contain over 124,000 discussion topics or threads, which are collections of posts displayed from oldest to most recent. Posts are user-submitted messages which include a set of data elements comprising user details (i.e., username, picture, and number of posts), the date and time of submission, and the content of the message. To date, CSN community discussion boards have over 1.2 million posts within their discussion threads. Figure 1 shows a screenshot of the CSN.

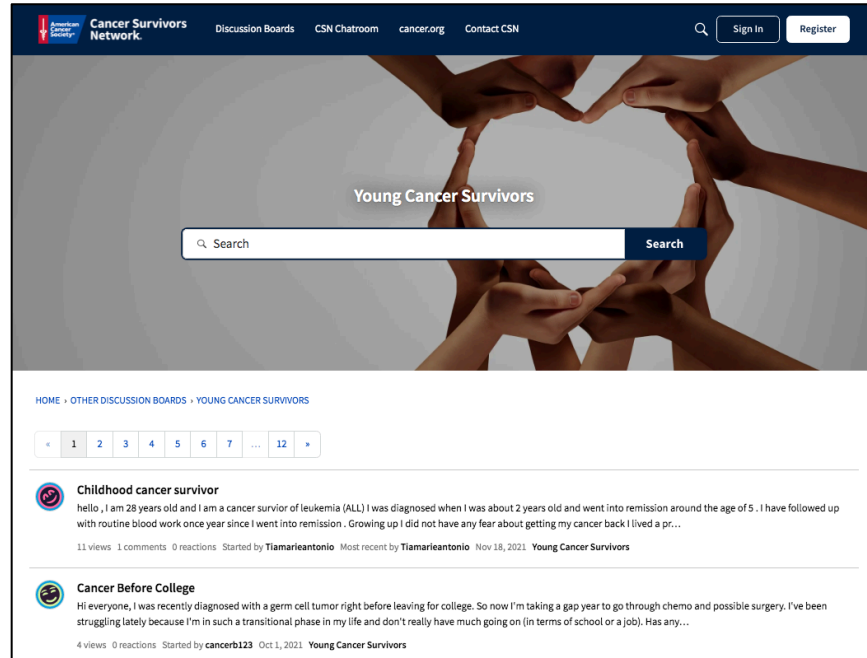


Figure 1. Screen capture of the CSN site.

Recent studies have shown that the CSN has over 3.6 million users from over 200 countries or territories (Fallon et al., 2018). In addition, discussion boards have been identified as the most used feature of the CSN (81.1%), followed by the search function and the member resource library (63.8% and 50.2%, respectively) (Fallon et al., 2018). Furthermore, Westmaas and colleagues (2019) identified the CSN’s most frequently used functions and classified them into three types of user engagement with the CSN community: social/communal, interpersonal communication, and informational/search (Westmaas et al., 2019). Social/communal engagement includes the frequent use of the CSN functions of “creating/writing one’s own blog,” “using the chat function,” and “adding friends.” Interpersonal communication engagement includes the frequent use of

the CSN functions of “reading discussions boards,” “posting to them,” and “read/send private messages.” Informational/search engagement includes the frequent use of the CSN functions of “using the CSN search function,” “reading blogs,” and “read/contribute to the member resource library” (Westmaas et al., 2019). Interestingly, while social/communal and interpersonal communication engagement types were associated with increased online social support, interpersonal communication showed a particularly strong association with online social support (Westmaas et al., 2019). Online social support exchanged through CSN was also associated with increased well-being among cancer survivors who reported low offline social support (Westmaas et al., 2019).

Despite the high demand for and potential benefits suggested in the literature, CSN remains an underexplored rich resource of online social support for cancer survivors. Research materials used for my study were extracted from the CSN. A dataset comprising all discussion topics (threads) from a CSN discussion board was used for this study and is described in the following chapter. CSN data were well suited for my study because the content of online interactions between CSN members allowed me to characterize social support, psychosocial needs, and linguistic features. In addition, CSN data allowed me to examine the structure of online interactions among AYA cancer survivors by focusing on a discussion forum specifically for this unique population. An integration of content and structure of online interactions between AYA cancer survivors provided me with the exceptional opportunity to contribute to the knowledge and understanding of supportive interactions taking place in an important online support resource for a special population with identified unmet psychosocial needs.

3.2 Data Collection and Sample

The CSN was selected for this study because of its relevance and levels of use, as described previously. Data for this study were extracted in July 2021 using Web scraping, a technique used for “extracting and combining contents of interest from the Web in a systematic way” (Glez-Peña et al., 2014, p. 789; Landers et al., 2016). The programming language Python was used to write a computer program to perform web scraping from the CSN (Python.org, n.d.). The Scrapy application framework was applied to extract data as HTML via the BeautifulSoup selector (Scrapy, n.d.). Web scraping was appropriate to extract behavioral data from CSN because it made the collection process “invisible” to CSN community participants, removing biases as well as risks associated with researcher contamination (Landers et al., 2016).

The sample of messages drawn for this study comprised all discussion threads (topics) from the CSN discussion board “Young Cancer Survivors” (to maintain continuity of interactions between and among members). A total of 329 threads containing 1,710 individual messages were collected for analysis ($n = 1,710$). Each thread included the original message post and all individual replies to the original message. All messages were posted by 747 CSN discussion board participants between March 2004 and July 2021. Given the nature of the topics discussed (hair loss, dating, college, etc.), it was assumed that members participating in this forum included AYA cancer survivors as well as their caregivers, family members, and friends. Figure 2 shows an overview of the CSN discussion board “Young Adult Survivors” patterns over the time period. Ninety percent (295) of topics received between 0 and 8 replies, indicating that most

communication interactions taking place in the discussion board did not extend beyond 8 messages.

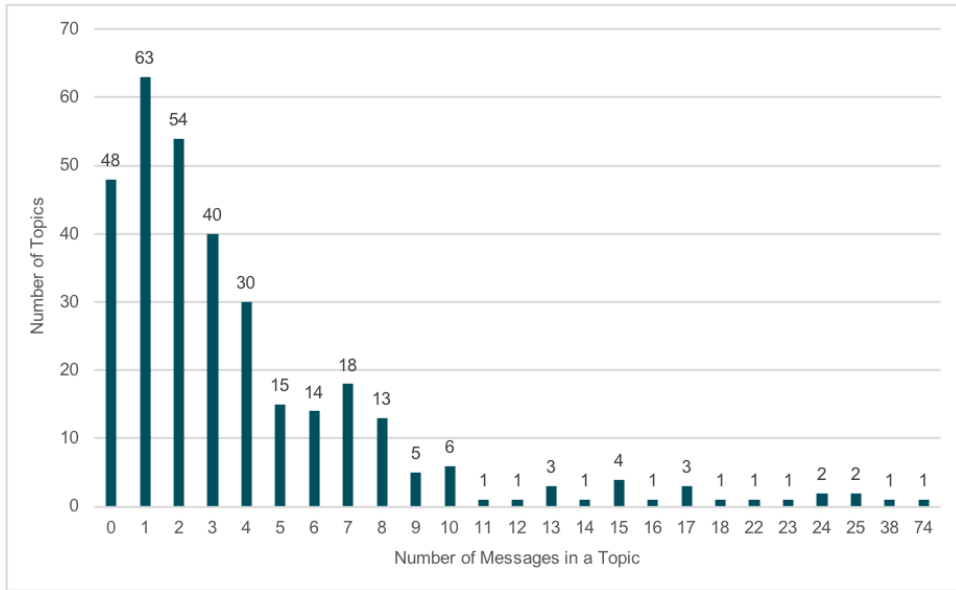


Figure 2. Frequency distribution of number of messages per topic.

In addition, Figure 3 shows CSN individual communication patterns over the time period. Sixty five percent (487) discussion board participants posted one message, 16% (121) participants posted two messages, and 15.9% (119) participants posted between three and eight messages, indicating that most discussion board participants posted between one and two messages.

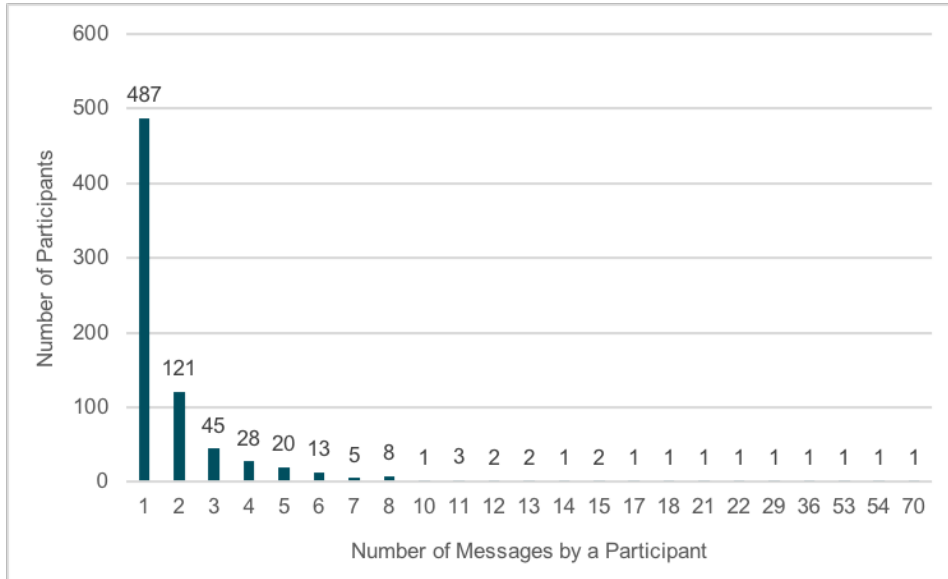


Figure 3. Frequency distribution of number of messages per participant.

**Chapter 4. Specific Aim 1: To Characterize the Content of Online Interactions
Among AYA Cancer Survivors in Terms of Types of Social Support, Survivorship
Care Plan Domains, Psychosocial Need Themes, and Linguistic Features**

The first step necessary to understanding the patterns of online relations is establishing the kinds and extent of social support, SCP domains, psychosocial need themes, and linguistic features observed in online interactions among AYA cancer survivors. This will help to clarify their impact on the exchange of social support targeting psychosocial needs. Once the presence and extent of such qualities are identified, it will become possible to examine how they are affected by patterns of online interactions between community participants.

Applying a content analysis approach allowed me to characterize the content of online interactions from the selected CSN discussion board. Content analysis is defined as “the systematic, objective, quantitative analysis of message characteristics” (Neuendorf, 2016). As such, content analysis is commonly used to “determine the presence of certain words, themes, or concepts within some given qualitative data (i.e., text)” (Columbia University, n.d.). Content analysis was used to achieve the first aim of this study because it could determine the extent to which online interactions among AYA cancer survivors contained social support, SCP domains, and psychosocial needs.

In addition, computerized linguistic analysis allowed me to extract the linguistic features embedded within online interactions, revealing how language is used by AYA cancer survivors to exchange social support online. To achieve Specific Aim 1, three

objectives were completed as presented next. Chapter 4 describes the methods and results for each objective of the first aim of the study.

4.1 Objective 1: Develop a Coding Scheme to Determine the Presence and Extent of Types of Social Support, Survivorship Care Plan Domains, and Psychosocial Needs in Online Interactions among AYA Cancer Survivors

Characterizing AYA cancer survivors' online interactions according to pre-established theoretical and empirical concepts started with the development of a sound coding scheme. The coding scheme served as the protocol for content-analyzing messages with the objective of minimizing individual coding differences between coders, thus reducing bias. The coding scheme consisted of the *codebook*, which contained all content analysis variables fully explained and operationalized, as well as the *coding form*, which provided the instrumental support for the coding of messages. These two components provided sufficient details to guide coders through the coding process.

Developing an objective and reliable coding scheme for this study required a careful process aligned with scientific content analysis. Figure 4 shows a flowchart of the content analysis process, with eight steps outlined. Step 1 (theory and rationale) was based on the literature review and provided a theoretical foundation as well as a rationale for the analyzed content. Steps 2 (conceptualizations) and 3 (operationalizations) focused on identifying the variables to be measured in the content analysis. These steps provided conceptual definitions for all variables and identified the units of data collection and their levels of measurement. With these elements, step 4 (coding scheme) resulted in an initial coding scheme that was pilot-tested and assessed for reliability using a subsample of

instances (steps 5 (sampling) and 6 (training and pilot reliability), respectively). Once the coding scheme was pilot-tested and determined to be reliable, it was used to manually code all remaining messages and report final results.

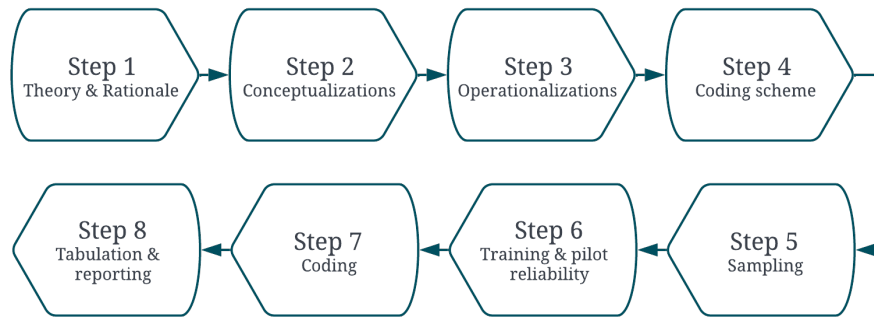


Figure 4. Content analysis process.

In this section, I describe the methods followed to develop the coding scheme used to characterize the content of online interactions among AYA cancer survivors. The procedures carried out for pilot testing and assessing the coding scheme's reliability are presented together with the final coding scheme.

4.1.1 Methods

4.1.1.1 Unit of Analysis. The unit of analysis for the content analysis was the individual message posting, defined as a single message shared (or posted) by an individual CSN discussion board participant. A single message contained from a single word to multiple paragraphs (Campbell-Eichhorn, 2008, Crook & Love, 2017).

4.1.1.2 Measures. Review of the literature identified key study variables. All variables used in the content analysis and their conceptual definitions were identified as

part of steps 2 and 3 of the content analysis process. These variables were grouped into four dimensions and are described next.

1. *Social Support.* Social support has been widely studied in the context of online communities and a number of coding schemes originally developed for face-to-face contexts have been adapted by researchers. The Social Support Behavior Code (SSBC; Cutrona & Suhr 1992, 1994; Suhr, 1990; Suhr et al., 2004) was adapted for use in this study because it has been successfully used in the past to study online support groups (Braithwaite et al., 1999; Coulson et al., 2007; Coulson & Greenwood, 2012; Donovan et al., 2014; Yu et al., 2017). The SSBC was originally developed to “assess social support behaviors in the context of help-intended dyadic interactions” over time (Suhr et al., 2004). Moreover, the SSBC assesses social support as a multidimensional construct, thus providing a typology of social support, which assesses 31 individual behaviors that fall into six mutually exclusive categories: informational support, emotional support, esteem support, tangible aid, network support, and negative behaviors (see Table 1; Cutrona & Suhr 1992, 1994; Suhr, 1990; Suhr et al., 2004). In addition, two categories were added to the coding scheme to distinguish whether social support was requested or provided in online interactions among AYA cancer survivors. Social support requested/provided was measured in line with Bambina’s (2007) method.

2. *Survivorship Care Plan Domains.* The cancer journey does not end with the completion of primary treatment (The Livestrong Foundation [Livestrong], 2011). This reality led the IOM to outline the core elements of survivorship care as well as recommendations aimed at enhancing the quality of care for cancer survivors (IOM,

2006). One of the cornerstones of survivorship care is the Survivorship Care Plan (SCP) (IOM, 2006). A SCP is the “record of a patient’s cancer history and recommendations for follow-up” (Houlihan, 2009) and its purpose is to “personalize and coordinate care by sharing the document with the survivors’ other healthcare providers” (Palos & Gilmore, 2019). Moreover, the core elements needed for any SCP were consensually identified by survivorship experts in 2011 (Livestrong, 2011). Given the importance of this communication tool, this study adapted the survivorship care domains as applied by an MD Anderson’s survivorship clinic (Palos & Gilmore, 2019; Rodriguez & Lewis-Patterson, 2019). A SCP template created and used by MD Anderson, developed in line with the IOM and the Livestrong Foundation’s survivorship core elements of survivorship care, provided the four SCP domains categories of: cancer surveillance and screening; late effects/side effects management; risk reduction and cancer detection; and psychosocial functioning (IOM, 2006; Livestrong, 2011; Palos & Gilmore, 2019; Rodriguez & Lewis-Patterson, 2019).

3. *Psychosocial Need Themes.* Cancer survivors and their caregivers experience multiple psychosocial challenges, as described in Chapter 2 (Janssen et al., 2021; Jin et al., 2021; Perez et al., 2020). In particular, AYA cancer survivors have multiple developmental and age-specific psychosocial needs that remain unmet (Janssen et al., 2021; Jin et al., 2021; Perez et al., 2020). In order to identify and analyze psychosocial issues from online interactions among AYA cancer survivors, this study adapted a taxonomy of psychosocial needs of cancer survivors developed by Burg and colleagues (Burg et al., 2015). Burg and her team identified 16 themes of unmet psychosocial needs

from a subsample of 1,514 national cross-sectional surveys of cancer survivors (Burg et al., 2015). Consequently, 16 categories were adapted to code psychosocial need themes: Physical, Financial, Education/Information, Personal Control, System-of-Care, Resources, Emotions/Mental Health, Social Support, Societal, Communication, Provider Relationship, Cure, Body Image, Survivor Identity, Employment, and Existential (Burg et al., 2015).

4. *Message Author.* A fourth dimension was added to the code scheme with the aim of identifying the author of each individual message posting. Given that the CSN is an OSC for cancer survivors and caregivers, it was assumed that message authors would include either cancer survivors or caregivers. However, the cancer survivor literature recognizes caregivers, family, and friends as survivors (Rodriguez & Lewis-Patterson, 2019). Accordingly, five categories were adapted: (1) cancer survivor (including individuals diagnosed, undergoing treatment, or in remission of cancer); (2) family caregiver; (3) informal caregiver; (4) other; and (5) unknown.

4.1.1.3 Coding Scheme Development Methodology. A preliminary coding scheme was created by integrating the four dimensions of social support, survivorship care plan domains, psychosocial needs, and message author. These dimensions contributed a preliminary set of content analysis variables that were fully defined and operationalized in the codebook. Additionally, a preliminary code form was generated from the codebook and imported into content analysis software to perform pilot testing. Next, the preliminary coding scheme was evaluated through pilot-test iterations carried out by a team of researchers as well as inter-rater reliability assessments on all study

variables to obtain a robust coding scheme. Such pilot-test iterations and resultant revisions provided materials for fine-tuning the preliminary coding scheme until the final version used in the content analysis was obtained. The development and evaluation of the coding scheme were carried out with Dedoose (version 9.0.17), a web-based application that allows the management, analysis, and presentation of qualitative and mixed method research data in a flexible, secure, and collaborative way (SocioCultural Research Consultants, LLC, 2021).

4.1.1.4 Coding Scheme Evaluation. The preliminary version of the coding scheme was pilot-tested and fine-tuned before the final version was achieved. Accordingly, validity was assessed to achieve reliability, thus ensuring the development of a valid, reliable, and useful code scheme.

Pilot Testing. A rigorous content analysis not only requires a carefully developed coding scheme, but also rigorous training to achieve intercoder reliability (Neuendorf, 2016). For this reason, step 6 of the content analysis (Training and Pilot Reliability) followed a careful process that consisted of five iterative training/revision sessions carried out between the team of coders. The team consisted of two doctoral students (including the study author) who coded the messages applying the code scheme and one faculty member who facilitated the resolution of discrepancies between coders. The initial coding scheme was introduced and discussed between coders during the first training session. All variables were described, and their definitions and measures were discussed. Next, during the second training session the coders coded a small random subsample of 10 messages together to practice and engage in a consensus-building

discussion around the coding scheme measures. After this session, an additional random subsample of 10 messages was coded independently by each coder and results were discussed between coders during a third training session. Training sessions iterated on discussing and revising the coding scheme measures and examples drawn from the dataset while considering (1) the applicability of the coding scheme, (2) recurring variables not captured by the coding scheme, and (3) areas of confusion or confounding of variables. After the first three training sessions were completed, the coders independently coded a random subsample of 10% of the messages that represented the variety of the dataset (Neuendorf, 2016). Once the coding of 10% of messages was completed, the coders reconvened to discuss discrepancies on 10 randomly selected messages. During the final two sessions, the faculty member facilitated the resolution of discrepancies between coders and remaining discrepancies were resolved by coders reaching mutual consensus. As a result, the coding scheme was continuously revised and fine-tuned prior to the final coding based on the feedback obtained from the training sessions.

Inter-Rater Reliability. Inter-rater reliability (IRR) was assessed on each study variable using the results from the pilot testing (subsample of 10% of messages). Moreover, reliability coefficients were calculated using simple agreement and Dwet's AC_1 statistic. Simple agreement is one of the most extensively used coefficients in the literature and it assesses the coders' agreement on precise values assigned to cases on a given variable (Neuendorf, 2016). It is normally calculated as the "percent agreement" (PA) as the number of agreements divided by the total number of cases ($PA = A/n$, where

A is the number of agreements between two coders and n is the total number of cases coded (Neuendorf, 2016). However, one concern is its “failure to account for potential chance agreement and the rigid requirement of the precise matching of coder’s scores”, resulting in an insufficient measure of IRR (Neuendorf, 2016). Consequently, Neuendorf (2012) recommends reporting agreement only when accompanied by chance-corrected coefficients (Neuendorf, 2016).

Accordingly, the AC_1 statistic was also used to assess reliability. Gwet’s AC_1 was designed to overcome the impact of the two “*kappa* paradoxes”, which has been shown to impact reliability outcomes of unbalanced or skewed distributions (Feinstein & Cicchetti, 1990; Gwet, 2008; Neuendorf, 2016). The AC_1 statistic ranges from .00 (no agreement) to 1.00 (perfect agreement) and is calculated with the following formula:

$$AC_1 = \frac{PA - PA_E}{1 - PA_E}$$

where PA is the observed proportion agreement and PA_E is the proportion agreement expected by chance (Gwet, 2008, Neuendorf, 2016). Both reliability coefficients were calculated for all variables with R (version 4.0.3) through the R packages irr (version 0.84.1) and irrCAC (version 1.0) (Gwet, 2019; Gamer et al., 2019; R Core Team, 2020). In addition, Altman’s (1991) benchmarking guideline, which is widely used in the literature, was used to interpret reliability results (Altman, 1991; see Appendix A).

Finally, the overall IRR coefficients revealed reliability levels ranging from moderate to very good. Appendix B shows the reliability coefficients for all study variables. As for the dimension of social support, consistent with Cutrona and Suhr (1992), reliability calculations are reported on the six social support categories (Cutrona

& Suhr 1992); thus, simple agreement for all six categories ranged between 74% and 99%, while Gwet's AC_I scores ranged between .56 and .95. In addition, simple agreement for social support requested and social support provided was 88% and 68%, respectively, while Gwet's AC_I scores were .80 and .40, respectively. As for the dimension of SCP domains, simple agreement for all four categories ranged between 72% and 99%, while Gwet's AC_I scores ranged between .45 and .99. For the dimension of psychosocial need themes, simple agreement for all 21 categories ranged between 78% and 100%, while Gwet's AC_I scores ranged between .60 and 1.00. Last, for the dimension of message author, simple agreement for all four categories ranged between 87% and 100%, while Gwet's AC_I scores ranged between .79 and 1.00.

4.1.1.5 Code Scheme Design Adaptations. Coding scheme pilot-test iterations and revisions to the preliminary version resulted in modifications to three dimensions. First, for the social support dimension, the category of network support together with its three subcategories (access, presence, and companions) were retained with the SSBC adapted for this study, in line with previous versions of the tool (Cutrona & Suhr 1992, 1994; Suhr, 1990). The most recent version of the SSBC published in the literature removed the category of network support from the SSBC, due to low levels of applicability and IRR within the context of marital partners (Suhr et al., 2004). However, previous research in contexts of online support communities have shown network support to be exchanged among members of such communities (Braithwaite et al., 1999; Coulson et al., 2007; Coulson & Greenwood, 2012; Donovan et al., 2014; Yu et al., 2017). Particularly, Donovan et al. (2014) found that 24% of online exchanges among young

adult cancer survivors responding to uncertainty contained network support (Donovan et al., 2014). Similarly, the subcategories of referral and encouragement from the categories of informational support and emotional support, respectively, were also retained from previous versions of the tool with the SSBC adapted for this study (Cutrona & Suhr 1992, 1994; Suhr, 1990). Moreover, the subcategory of anchorage was added to the esteem support category in line with Coulson et al. (2007). Furthermore, the subcategory of physical affection from the emotional support category was modified to virtual affection to adapt it to an online context (Coulson et al., 2007). Finally, the subcategory of listening from the emotional support category was deleted because of the impossibility to code it within online contexts such as online communities (Braithwaite et al., 1999; Coulson et al., 2007).

Second, for the psychosocial need themes dimension, researchers agreed after pilot testing the code scheme to add five new categories of themes that were being discussed among AYA cancer survivors. Because of the relevance and age-appropriateness of the themes, the following categories were added: peer support, school, dating and relationships, fertility/reproductive health, and parenthood. Last, for the message author dimension, researchers agreed after pilot testing the code scheme to add the two new categories of other: CSN staff/moderator and other: undiagnosed patient. In addition to allowing the coding of messages created by CSN staff/moderators, researchers thought it was relevant to differentiate those participants who engaged with the online community without being formally diagnosed with cancer or being a cancer survivor.

4.1.1.6 Ethical and Legal Considerations. This study focuses on online social support exchanges among CSN members while interacting via one of CSN's discussion boards. Discussion boards or forums constitute a mode of synchronous computer-mediated communication (CMC) facilitated by the Internet. Online research and the ways Internet users act and interact via diverse CMC modes requires careful attention to the protection of human subjects and related matters (Jacobson, 2010). Despite CSN's nature of being publicly open to any cancer survivor using the Internet, research using data from this online community required taking into consideration the protection of human subjects, including matters of identity, anonymity, pseudonymity, consent, and privacy (Jacobson, 2010). Copyright matters were likewise considered (Jacobson, 2010).

Protection of Human Subjects. This study entailed the "observation of public behavior" because it examined CMC interactions among members of a CSN publicly available forum. Researchers agree that "cyberspace research entails 'human subjects'" (Jacobson, 2010) which requires a careful consideration when protecting human subjects given the levels of complexity that are introduced by the Internet when carrying out online research. In addition, despite the public availability of the data used in my study, human subjects regulations must be considered even when the data were extracted and recorded in a way that the identity of subjects cannot be determined. Furthermore, evidence shows that online research introduces additional challenges to protecting the identity of human subjects (Jacobson, 2010). Studies have shown that members in online communities may voluntarily disclose identifiable information, while some members

attempt to uncover the identities of other members by aggregating information shared within a given community or even across several communities (Jacobson, 2010).

This study used digital trace data of online interactions from all members participating in one CSN publicly available forum. Such data included members' usernames, text of messages posted by members (including initial threads and replies), and dates of posts. Despite CSN encouragements for users to protect their privacy by remaining anonymous and controlling their personal identifiers (i.e. usernames), risks for CSN members' **confidentiality** remain due to the issues mentioned above. This study protected members' confidentiality by de-identifying all identifiable data from CSN members of the forum being examined once the data extraction process took place. First, despite members' usernames being pseudonyms (thus anonymous), they were de-identified by assigning random user id's and repopulating the dataset, thus further protecting members' identity. Second, the Amazon Comprehend Medical service was used to further mitigate confidentiality risks by identifying personal health information (PHI). ACM is a natural language processing tool that uses machine learning to extract relevant medical information from unstructured text (Amazon Comprehend Medical, n.d.). Last, further identifiable information that may have been disclosed by members while interacting in the CSN forum being examined and which was not masked by the ACM tool, was replaced with the string "*****" by researchers during the manual coding of messages as part of the study's content analysis. The Safe Harbor de-identification standard was used to guide the manual de-identification of forum messages (IOM, 2015). The Safe Harbor standard includes 18 data identifiers, including: all geographic

subdivisions smaller than a state, including street address, city, etc.; all elements of dates (except year) for dates directly related to an individual, including birth date, etc.; telephone numbers; electronic mail addresses; social security numbers; medical record numbers; certificate/license numbers; and any other unique identifying number, characteristic, or code; among others (IOM, 2015; U.S. Department of Health and Human Services, [HHS], 2012).

Privacy. Privacy is another ethical issue concerning online research. Jacobson (2010) argues that “[c]hat rooms and virtual communities are public places in the sense that anyone with access to the Internet may connect to them, either as a temporary visitor or as a registered member” (Jacobson, 2010). In this study, given that all CSN forums are public, messages posted to any forum are not considered private and therefore are not protected by privacy laws. Finally, considering the character of this study as a naturalistic observation of online public behavior as well as the public nature of CSN forums, the risk of harm to subjects that could result from a breach of confidentiality is minimal; thus, analyses required for this study were conducted without obtaining informed consent from subjects.

Legal Considerations. On the one hand, because messages are posted on publicly available forums within online communities, such as the CSN, they are not protected by human subject regulations or by privacy laws. On the other hand, asynchronous CMC, including Internet forums, are protected by copyright laws (Jacobson, 2010). However, the doctrine of fair use allows researchers to use, without the consent of its owner, otherwise copyrighted materials (Jacobson, 2010). Fair uses are limited to “purposes such

as criticism, comment, news reporting, teaching, scholarship, or research,” (Jacobson, 2010, p. 138) suggesting that researchers are free from copyright restrictions (Jacobson, 2010).

Institutional Review Board (IRB). The research protocol of this study was submitted to UTHHealth’s IRB for review before data collection and received the status of exempt (IRB letter number HSC-SBMI-21-0237 of March 23, 2021).

4.1.2 Results

The final coding scheme was theoretically and empirically influenced by diverse sources. Following a scientifically grounded process to carry out the content analysis resulted in a concise coding scheme that extends previous studies of online social support by including survivorship care plan domains as well as psychosocial need themes. In addition, IRR assessments showed good confidence and validity for all study variables, making the coding scheme a trustworthy and valid coding tool. An overview of the final coding scheme with all dimensions, categories and subcategories is shown in Table 1. A complete code scheme including all categories and subcategories with definitions and examples drawn from the dataset to illustrate them is presented in Appendix C.

Table 1

Coding scheme overview

Social Support	
Category	Subcategory
Informational Support	Suggestion and advice
	Referral
	Situation appraisal
	Teaching

Category	Subcategory
Emotional Support	Relationship Virtual affection Confidentiality Sympathy Understanding and empathy Encouragement Prayer Expresses concern Reassurance
Esteem Support	Compliment Validation Relief of blame Anchorage
Tangible Aid	Loan Direct task Indirect task Active participation Willingness Complies with request
Network support	Access Presence Companions
Negative Behaviors	Interrupt Complain Criticism Isolation Disagree or disapprove
Social Support Requested	
Social Support Provided	
Survivorship Care Plan Domains	
Cancer surveillance and screening	
Late effects/side effects management	
Risk reduction and cancer prevention	
Psychosocial functioning	
Psychosocial Need Themes	
Physical	Cure
Financial	Body image
Education and information	Survivor identity
Personal control	Employment
System-of-care	Existential/spiritual
Resources	Peer support*
Emotional and mental health	School*
Social support	Dating and relationships*

Psychosocial Need Themes	
Societal	Fertility/reproductive health*
Communication	Parenthood*
Provider relationship	
Message Author	
Cancer survivor	Unknown
Family caregiver	Other: CSN staff/moderator*
Informal caregiver	Other: Undiagnosed patient*
Other	

* Categories that were added after the pilot test of the coding scheme.

The coding scheme contains six categories of social support: (1) informational support (providing information about the problem, how to appraise the problem, or how to cope with the problem); (2) emotional support (communicating empathy or caring); (3) esteem support (communicating confidence in the other's worth, competence, or ability to solve the problem); (4) tangible aid (offering assistance or tangible resources to help solve the problem); (5) network support (communicating belonging to a group of persons with a similar interests and concerns); and negative behaviors (sarcasm, criticism, disagreement, interruption, complaint, refusals to help). These six social support categories comprise a total of 31 subcategories retained from Cutrona's SSBC (Cutrona & Suhr 1992, 1994; Suhr, 1990; Suhr et al., 2004). In addition, two social support categories were added to code whether a message provided or requested social support (Bambina, 2007). Moreover, the coding scheme contains four SCP domain categories that were retained from an SCP template developed by MD Anderson: (1) cancer surveillance and screening (detection and treatment of late malignancy recurrence or new second malignancies); (2) late effects/side effects management (health maintenance and observation of vital organ function); (3) risk reduction and cancer detection (lifestyle

changes to prevent cancer and risk assessment); and (4) psychosocial functioning (psychosocial support services to maintain healthy relationships and restored life) (IOM, 2006; Livestrong, 2011; Palos & Gilmore, 2019; Rodriguez & Lewis-Patterson, 2019). Furthermore, the coding scheme contains 21 categories of psychosocial need themes. Sixteen categories were retained from a taxonomy developed by Berg et al (2016) while five new categories were added by researchers after pilot testing the coding scheme (Berg, 2016). Lastly, the coding scheme contains seven message author categories that were added to code the author of the messages.

4.1.3 Conclusion

The final coding scheme for this study is a concise, trustworthy, and valid content analysis tool that extends previous studies of online social support by including Survivorship Care Plan domains as well as a comprehensive taxonomy of psychosocial need themes.

4.2 Objective 2: Determine the Presence and Extent of Types of Social Support, Survivorship Care Plan Domains, and Psychosocial Need Themes of Online Interactions Among AYA Cancer Survivors

Equipped with a robust coding scheme, the final steps of the content analysis process consisted of coding all remaining messages (step 7) and tabulating and reporting findings (step 8). In this section, I describe the final results of the content analysis after all remaining messages were coded by applying the developed coding scheme with the software tool Dedoose.

4.2.1 Methods

4.2.1.1 Data Analysis. Descriptive statistics were used to report the results of all variables. Frequencies of occurrence and percentages for each code category were calculated to analyze the types of support exchanges and psychosocial needs among members of the CSN discussion board as well as the overall levels of supportive interactions (i.e., number of codes per posts). In addition, bar graphs were used to visualize the frequencies and percentages of types of social support exchanges as well as psychosocial needs.

4.2.2 Results

All 1,710 individual message postings were coded using the final coding scheme. A total of 14,032 codes were manually applied and final results of the content analysis are presented next by coding dimension.

4.2.2.1 Social Support. Table 2 shows the frequency for each social support category along with the frequencies of social support requested and provided. The most frequently coded category was emotional support (60%), followed by informational support (38%), esteem support (32%), and network support (17%), with tangible aid (6%) and negative behaviors (1%) coded much less frequently.

Table 2

Descriptive statistics for messages containing social support categories

Social Support Category	Frequency	Percent of total messages
Emotional Support	1032	60.4

Social Support Category	Frequency	Percent of total messages
Requested	472	
Provided	751	
Informational Support	659	38.5
Requested	275	
Provided	469	
Esteem Support	554	32.4
Requested	207	
Provided	472	
Network support	296	17.3
Requested	158	
Provided	182	
Tangible Aid	111	6.5
Requested	46	
Provided	78	
Negative Behaviors	23	1.3
Requested	0	
Provided	23	

In addition, Figure 5 shows the number of messages that either requested or provided social support within each category. Similarly, out of a total of 613 support requests and 1,072 support provisions coded, the number of messages kept the same

trend, with emotional support being the support category most requested (28%) and provided (44%), followed by informational support (16% and 27%, respectively), esteem support (12% and 28%, respectively), and network support (9% and 11%, respectively). However, Figure 5 also shows that for each social support category, the amount of support provided exceeded the demand.

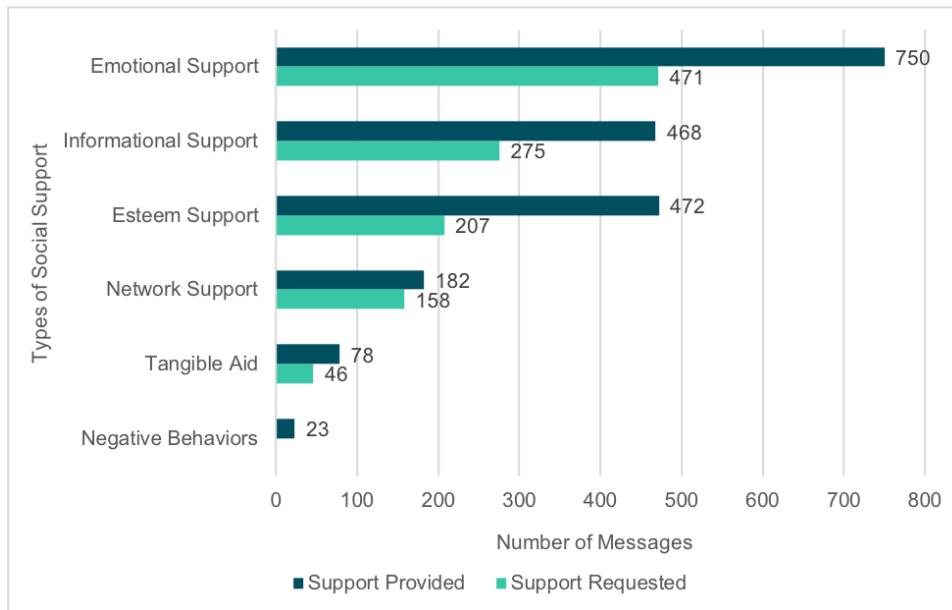


Figure 5. Number of messages by types of social support requested and provided.

4.2.2.2 Survivorship Care Plan Domains. A total of 1,553 (91%) messages were coded with SCP domains. Table 3 shows the number of messages coded within each SCP domain category. The majority of messages contained psychosocial functioning (80%), while the remaining three SCP domains categories were present far less often: cancer surveillance and screening (2.7%), late effects/side effects management (7.9%), and risk reduction and cancer prevention (0.5%).

Table 3

Descriptive statistics for messages containing SCP domains categories

SCP Domain	Number of Messages
Cancer surveillance and screening	47
Late effects/side effects management	135
Risk reduction and cancer prevention	8
Psychosocial functioning	1363

In addition, Figure 6 shows the percentage of messages that either requested or provided social support within SCP domains categories. While the majority of messages that requested or provided social support was aligned with psychosocial functioning (84% and 89%, respectively), messages that were aligned with cancer surveillance and screening and late effects/side effects management that provided support did not match the demand for social support.

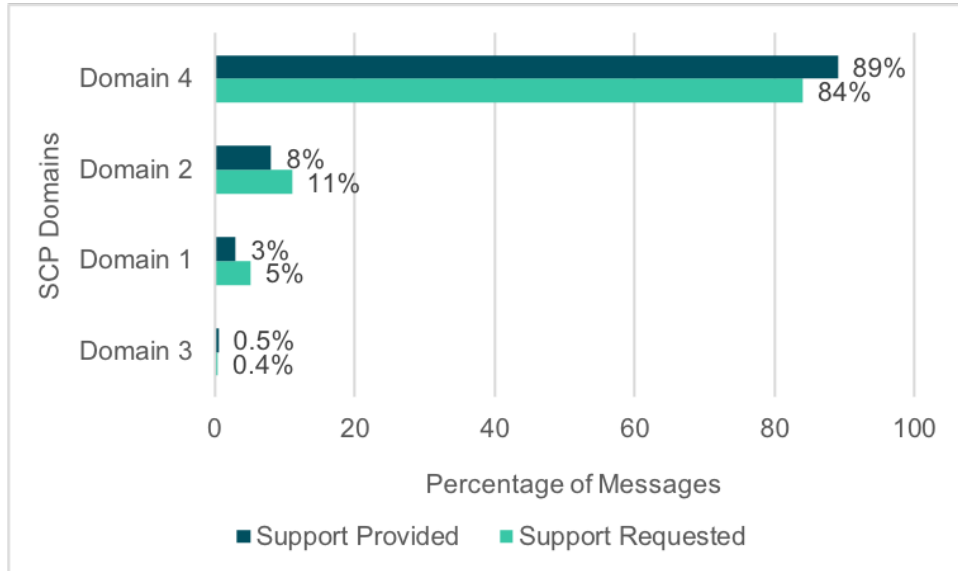


Figure 6. Proportions of messages that requested and provided social support aligned with SCP domain categories.

Lastly, Figure 7 shows the number of messages that contained social support aligned with SCP domain categories. The majority of support messages that aligned with the SCP domains categories of psychosocial functioning (domain 4) and late effects/side effects management (domain 2) contained emotional support (923 and 116, respectively), informational support (608 and 74, respectively), and esteem support (506 and 74, respectively).

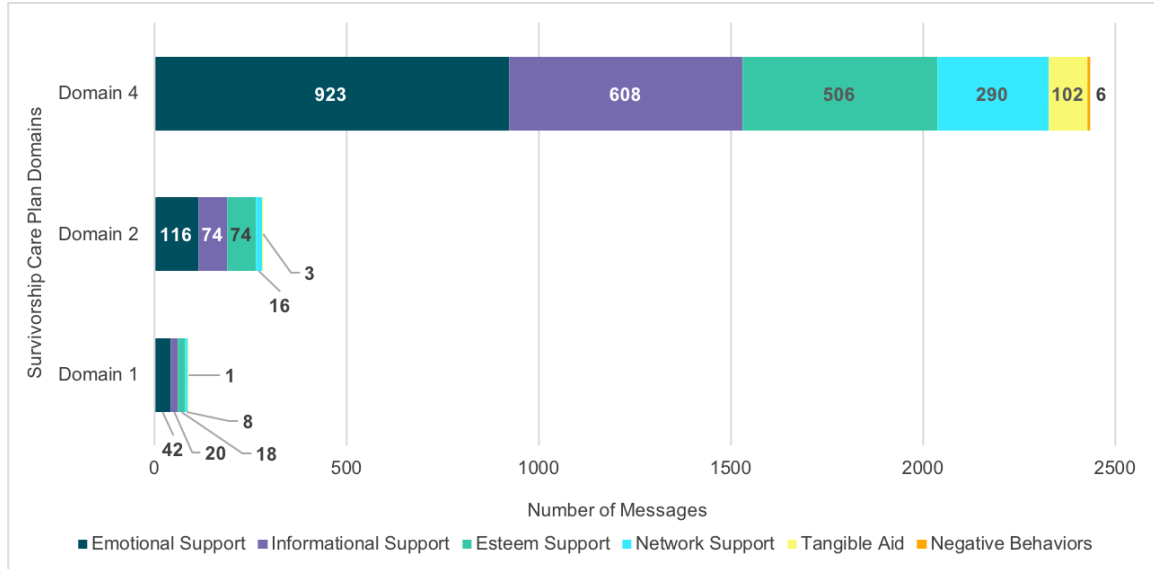


Figure 7. Number of messages that contained social support aligned with SCP domain categories (Domain 1: cancer surveillance and screening; Domain 2: late effects/side effects management; Domain 4: psychosocial functioning).

4.2.2.3 Psychosocial Need Themes. A total of 2,601 individual message postings contained psychosocial need themes. Figure 8 shows the frequency of psychosocial need themes. Six categories accounted for the majority of themes (74%): social support (23.4%), peer support (15%), emotions/mental health (14.3%), dating and relationships (10.3%), physical (5.7%), and education/information (5%). These categories were followed by body image (3.7%), fertility and reproductive health (3.5%), communication (3.2%), parenthood (3%), personal control (2.8%), school (2.2%), financial (1.5%), existential (1.4%), societal (1.4%), system of care (1%), resources (1%),

provider relationship (0.6%), employment (0.3%), survivor identity (0.3%), and cure (0.2%).

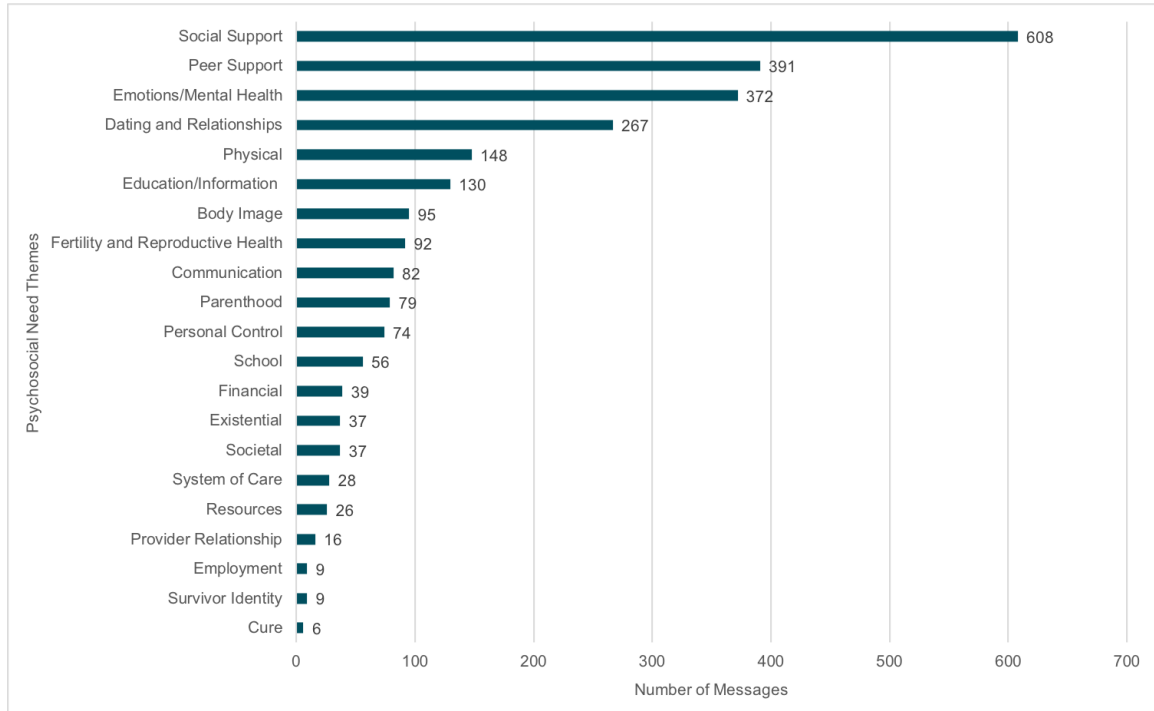


Figure 8. Frequency of psychosocial need themes.

In addition, Figure 9 shows the distribution of the six most frequent psychosocial need themes by types of social support. For emotional support, social support was the most frequent psychosocial need theme, followed by emotions/mental health, peer support, dating and relationships, and physical. For informational support, emotions/mental health was the most frequent psychosocial need theme, followed by social support, dating and relationships, education/information, peer support, and physical. For esteem support, emotions/mental health was the most frequent psychosocial

need theme, followed by social support, dating and relationships, peer support and physical. For network support, social support and peer support were the most frequent psychosocial need themes. Likewise, for tangible aid, social support and peer support were the most frequent psychosocial need themes.

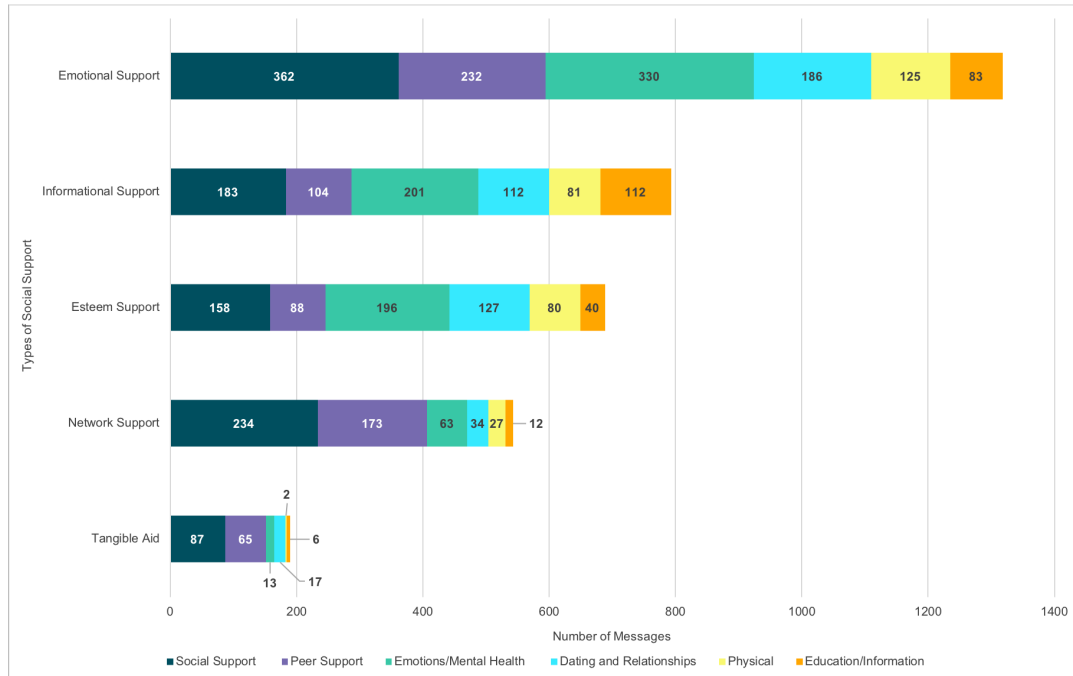


Figure 9. Frequency of psychosocial need themes by types of social support.

Finally, Figure 10 shows the distribution of the six most frequent psychosocial need themes by SCP domains. Overall, emotions/mental health, social support, and peer support were the most frequently discussed psychosocial need themes within SCP domains. As for SCP domain 4, social support was the most frequent psychosocial need theme, followed by peer support, emotions/mental health, dating and relationships,

physical, and education/information. As for SCP domain 2, physical and emotions/mental health were the most frequent psychosocial need themes.

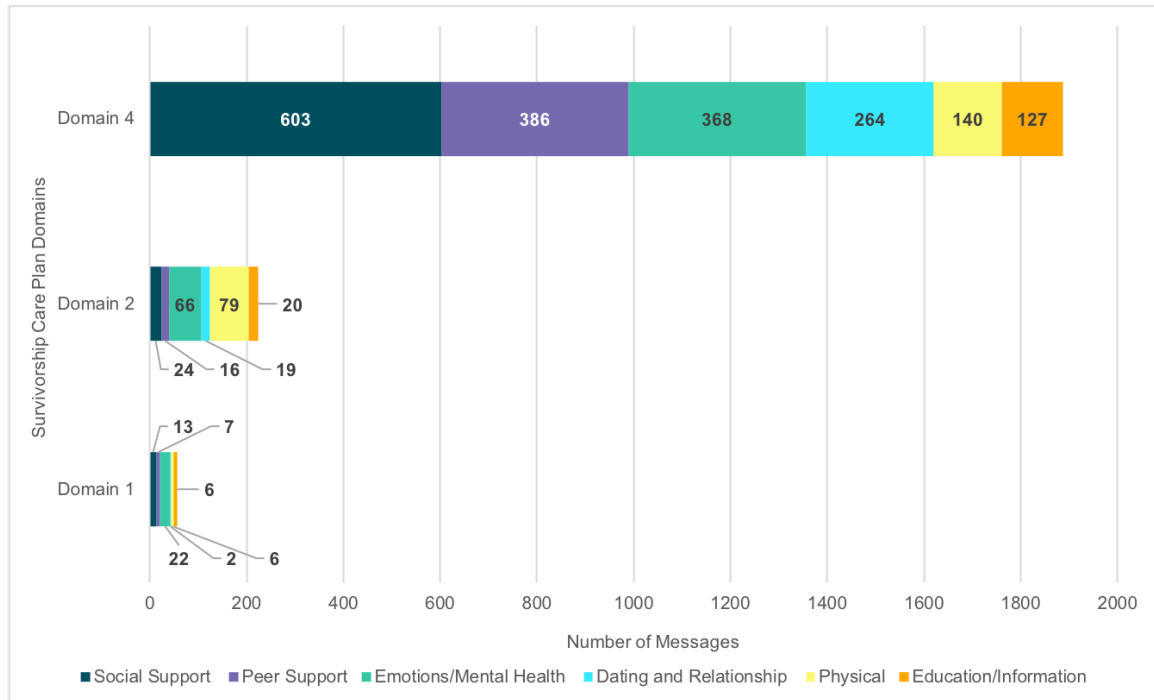


Figure 10. Frequency of psychosocial need themes within SCP domains (Domain 1: cancer surveillance and screening; Domain 2: late effects/side effects management; Domain 4: psychosocial functioning).

4.2.2.4 Message Author. Table 4 shows the frequency of message author categories. The majority of individual message postings were made by cancer survivors (82%) and family caregivers (4%). In addition, a small proportion of messages were made by informal caregivers (1%), CSN staff/moderators (1.6%), undiagnosed patients

(1%), and others (1%). Finally, a small proportion of messages whose authors could not be determined was coded as unknown author (5.7%).

Table 4

Descriptive statistics for message author categories

Message Author	Number of Messages	Percentage
Cancer Survivor	1413	82.6
Family Caregiver	72	4.2
Informal Caregiver	14	1
Other: CSN Staff/Moderator	28	1.6
Other: Undiagnosed Patient	20	1.2
Other	18	1
Unknown Author	98	5.7

4.2.3 Conclusions

The content analysis found emotional, informational, and esteem support types were the most requested among AYA cancer survivors. In addition, the results show that network support was also requested, while tangible aid was far less requested and negative behaviors were rare among AYA cancer survivors. Moreover, the data show social support provided far exceeded the requests for emotional, informational, and esteem support, suggesting that participants received the type of support they requested. Therefore, the CSN community appears to represent a valuable source of social support.

As for SCP domains, the results show that most online exchanges among AYA cancer survivors were related to SCP domains. Particularly, the majority of online exchanges were related to the SCP domain of psychosocial functioning, with the SCP domains of late effects/side effects management and cancer surveillance and screening observed far less frequently. In addition, the data show that while social support provided matched the amount requested within the SCP domain of psychosocial functioning, it did not match the amount of support requested within the SCP domains of cancer surveillance and late effects/side effects management. As for psychosocial need themes, the results show that a great variety of psychosocial needs are being discussed online among AYA cancer survivors. The majority of messages contained psychosocial need themes of social support, peer support, emotions/mental health, dating and relationships, physical, and education/information. Moreover, the data show that online social support is being exchanged concerning all psychosocial need themes. For example, informational support exchanges were related to dating and relationships and education/information; esteem support exchanges were related to emotions/mental health and dating and relationships; and network support exchanges were related to social support and peer support. Furthermore, the data also show that psychosocial need themes are likewise aligned with SCP domains. For example, the majority of psychosocial need themes that aligned with the SCP domain of late effects/side effects management included physical and emotions/mental health. Lastly, the results showed that the majority of messages were posted by cancer survivors and family caregivers.

4.3 Objective 3: Extract the Linguistic Features of Supportive Online Interactions Among AYA Cancer Survivors

Research shows the words individuals use to communicate reflect their psychological states. Consequently, examining the words individuals use in daily life allow researchers to identify their focus of attention, providing insights into their thoughts, what they are trying to avoid, their feelings, and the ways they are organizing and analyzing their worlds (Tausczik & Pennebaker, 2010). OSCs such as the CSN provide the opportunity to study the language cancer survivors use to exchange social support in a naturally occurring online environment. Accordingly, I used computerized text analysis to explore how AYA cancer survivors are using language to exchange social support within the CSN online support community. The linguistic features from the content of online interactions from the CSN discussion board “Young Cancer Survivors” were extracted using a well-established text analysis software. Online support exchanges were examined, and psychologically relevant categories were extracted with LIWC to better understand how AYA cancer survivors used language when providing and requesting online support. Subsequently, linguistic features were assessed for differences in language use between the groups of support providing and support requesting exchanges. In this section, I describe the methods used to extract the linguistic features from the content of online support interactions among AYA cancer survivors. Results from the linguistic analysis are presented together with their differences between groups of support providing and support requesting exchanges.

4.3.1 Methods

4.3.1.1 Computerized Text Analysis. Computerized text analysis has long been a central method to psychological language analysis, enabling the extraction of psychological insights from an individual's verbal behavioral data (Boyd & Schwartz, 2021). Broadly speaking, computerized text analysis aims to operationalize, measure, and quantify human communication, allowing researchers to extract diverse content features, including linguistic and stylistic features. Linguistic features include word classes such as nouns, verbs, pronouns, etc. as well as punctuation and emoticons (in social media-related communications). Lexical features include markers for semantic concepts or psychological constructs including cognitive processes or emotionality. In addition, computerized text analysis is commonly used to extract frequent themes or topics from a given text. Hence, computerized text analysis has been widely used as a top-down (or theory driven) approach to text analysis, with predefined theoretical categories (Moessner et al., 2018).

Computerized text analysis tools such as LIWC are based on the principle of “words frequencies represent attentional habits” (Boyd & Schwartz, 2021; Pennebaker et al., 2015; Tausczik & Pennebaker, 2010). By determining word frequencies, computerized text analysis tools enable researchers to measure microscopic language differences that lead to individuals' attentional focus, thus opening a direct channel into their mind. Equipped with such tools, researchers have determined that small particles of language such as pronouns and negations can provide valuable insights into diverse psychological phenomena. In addition, these verbal behavioral markers have been shown

to be reliable indicators of how individuals think, regardless of what they talk about (Boyd & Schwartz, 2021; Tausczik & Pennebaker, 2010). For instance, studies show that individuals who use high rates of articles and prepositions tend to focus on formal or concrete concepts and their inter-relations. Moreover, individuals with social status and confidence tend to focus on the external social environment versus themselves, as reflected by higher uses of “you” and “we” words than “I” words (Boyd & Schwartz, 2021; Tausczik & Pennebaker, 2010). Furthermore, insights into individuals’ personalities, life experiences, cultures, and societies can be gathered (Boyd & Schwartz, 2021; Tausczik & Pennebaker, 2010). Given that the “words as attention” theoretical paradigm allows researchers to establish the associations between cognitive words with things like stress and problem-solving, for example, an individual’s stress level can be similarly indirectly estimated (Boyd & Schwartz, 2021).

LIWC. The Linguistic Inquiry and Word Count (LIWC, version 2015) was used to extract the linguistic features from the content of online interactions from the CSN discussion board “Young Cancer Survivors” (Pennebaker et al., 2015). LIWC is a well-established text analysis software that reveals meaning along diverse psychological dimensions such as emotionality and social relationships by counting the words in such categories (Tausczik & Pennebaker, 2010). LIWC performs text analysis based on a robust internal default dictionary which guides its internal word-counting processes. LIWC’s default dictionary includes almost 6,400 words, word stems, and select emoticons (Pennebaker et al., 2015). In addition, LIWC distinguishes between the two broad word categories of *content* and *function* words. On the one hand, content words

include nouns, regular verbs, and many adjectives and adverbs, which convey the content of a communication (what individuals are saying). On the other hand, function words (or style words) include pronouns, prepositions, articles, conjunctions, auxiliary verbs, and a few other “esoteric categories”, which reflect how people are communicating (Tausczik & Pennebaker, 2010). The literature shows LIWC has been applied to study diverse psychological processes taking place online, such as the content from interactions within online support communities (Kornfield et al., 2018; Lewallen et al., 2014; Li et al., 2018; O’Dea et al., 2017; Xu & Zhang, 2016). As for AYA cancer survivors, the literature suggests that a deeper understanding of how this subpopulation uses language online can help them to exchange online support more effectively, for example, by writing messages that prompt social support exchanges (Crook et al., 2016). However, little is known about how AYA cancer survivors use language to interact with peers online and, more importantly, exchange social support. The literature review showed only four studies have attempted to examine the linguistic features of support exchanges among AYA cancer survivors within online communities (Crook et al., 2016; Kaal et al., 2018; Thompson et al., 2016; Warner et al., 2018). Findings from these studies were discussed in Chapter 2.

LIWC Dimensions. When LIWC performs text analysis, it calculates the percentage of total words by comparing all words against its internal dictionary. Once all text has been processed, LIWC provides close to 90 output variables, most of which are percentages of total words that matched its internal dictionary (Pennebaker et al., 2015). There are six exceptions: word count (WC), mean words per sentence (WPS), and four

summary variables: Analytic, Clout, Authentic, and Tone. WC and WPS provide raw word count and average words per sentence, respectively, while the summary variables provide percentiles based on standardized scores that are based on large comparison samples obtained from previous research. Thus, LIWC's output variables consist of 21 standard linguistic dimensions, 41-word categories focusing on psychological constructs, 6 personal concern categories, among others (Pennebaker et al., 2015). Appendix D shows a complete list of LIWC categories.

4.3.1.2 Data Analysis. Descriptive statistics were used to report the results of relevant LIWC output categories. First, the mean ratios and frequencies of occurrence for each LIWC output variable were calculated to analyze the overall linguistic features. In addition, mean ratios and frequencies of occurrence were calculated by social support provided and requested categories to analyze the overall linguistic features of support providing and requesting exchanges. Furthermore, bar graphs were used to visualize LIWC mean ratios across social support categories and histograms were used to visualize the distributions of word counts and LIWC output variable categories. Second, LIWC output variables were assessed for significant associations and differences. Specifically, based on the literature described in the background section, the following hypothesis guided the analysis:

- *H1*: There are no language use differences between online exchanges that requested social support and online exchanges that provided social support.

To assess the hypothesis, independent samples *t* tests were performed to assess whether mean word ratios differed significantly for AYAs online support requesting exchanges

compared to support providing exchanges. Preliminary data analysis indicated that LIWC output scores were multimodal, but the departure from normality did not require the use of a nonparametric test. In addition, the assumption of homogeneity of variance was assessed by the Levene test, which indicated no significant violation of the equal variance assumption; therefore, the pooled variances version of the t test was used. Independent samples t tests were calculated with R (version 4.0.3) (R Core Team, 2020).

4.3.2 Results

The sample of 1,710 individual message postings comprised a total of 199,817 words. All messages were analyzed using LIWC and ratios or percentages of words were obtained for its linguistic variables (except for total word count and words per sentence categories), as well as the four summary variables, which are calculated differently, as described in the previous section. Mean word ratios were used to account for individual differences between AYA cancer survivors' online exchanges requesting social support ($N=613$) and online exchanges providing social support ($N=1,072$). Mean word ratios for all linguistic variables are provided in Appendix E. In addition, regarding the hypothesis proposed, results from independent samples t tests indicated significant differences between AYA cancer survivors' online exchanges that requested and provided social support with respect to language use. Appendix E also provides all t test results while the most relevant findings are presented next.

Online support requesting exchanges ($M=147.04$, $SD=136.78$) had higher mean word count per unit of analysis than online support providing exchanges ($M=129.65$, $SD=134.13$), $t(1683)=2.54$, $p<.05$, as well as a higher mean ratio of words longer than six

letters ($M=13.82$, $SD=4.89$) than support providing exchanges ($M=13.14$, $SD=5.73$), $t(1683)=2.46$, $p<.05$. As for LIWC summary measures, support requesting exchanges had a higher mean score of authenticity ($M=73.49$, $SD=29.03$) than support providing exchanges ($M=64.1$, $SD=32.13$), $t(1683)=5.97$, $p<.001$, while support providing exchanges had a higher mean score of emotional tone ($M=63.8$, $SD=33.06$) than support requesting exchanges ($M=48.83$, $SD=33.42$), $t(1683)=-8.90$, $p<.001$. Figure 11 shows the mean scores for LIWC summary measures for AYA cancer survivors' online exchanges that requested and provided social support.

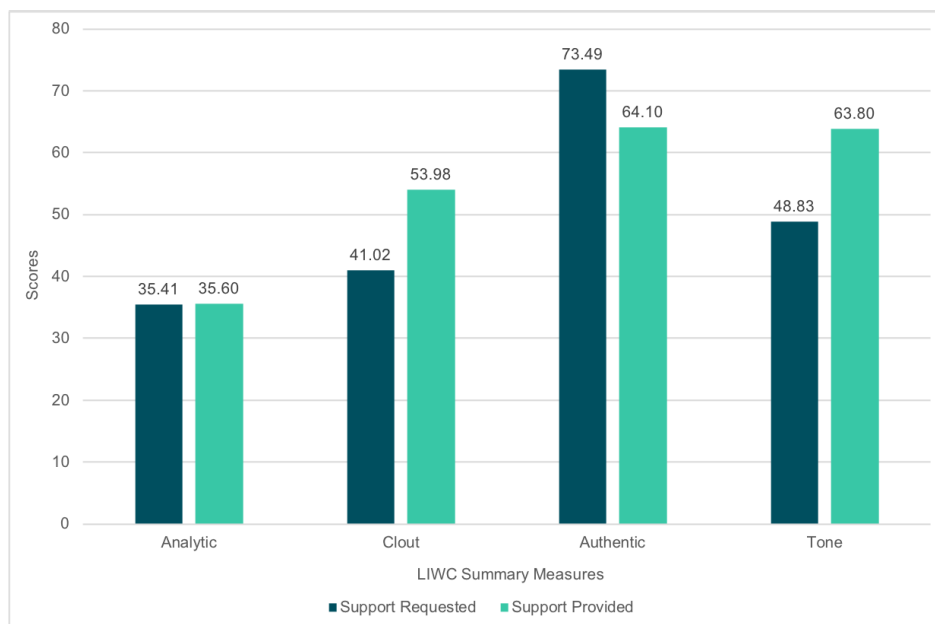


Figure 11. Mean scores for LIWC summary measures for AYA cancer survivors' online exchanges that requested and provided social support.

As for the linguistic dimension pronouns, support requesting exchanges had a higher mean ratio of first-person singular pronouns ($M=9.08$, $SD=3.9$) than support

providing exchanges ($M=7.84$, $SD=4.45$), $t(1683)=5.74$, $p<.001$, as well as a higher mean ratio of third-person singular pronouns ($M=0.75$, $SD=1.89$) than support providing exchanges ($M=0.51$, $SD=1.41$), $t(1683)=2.98$, $p<.01$. Moreover, support providing exchanges had a higher mean ratio of first-person plural pronouns ($M=0.63$, $SD=1.39$) than support requesting exchanges ($M=0.48$, $SD=1.14$), $t(1683)=-2.27$, $p<.05$, as well as a higher mean ratio of second-person pronouns ($M=3.36$, $SD=3.42$) than support requesting exchanges ($M=1.51$, $SD=2.29$), $t(1683)=-11.95$, $p<.001$, and a higher mean ratio of third-person plural pronouns ($M=0.58$, $SD=1.09$) than support requesting exchanges ($M=0.45$, $SD=0.86$), $t(1683)=-2.49$, $p<.05$.

Next, the LIWC categories for psychological processes with the most relevant findings included affective, social, and cognitive processes as well as drives. Figure 12 shows the mean percentage of total words for LIWC psychological categories of affective and social processes by exchanges that requested and provided social support. AYA cancer survivors used a higher number of words reflecting affective and social processes, particularly positive and negative emotions. As for affective processes, support requesting exchanges had a higher mean ratio of negative emotion words ($M=1.87$, $SD=1.82$) than support providing exchanges ($M=1.43$, $SD=1.85$), $t(1683)=4.72$, $p<.001$, while support providing exchanges had a higher mean ratio of positive emotion words ($M=4.68$, $SD=4.09$) than support requesting exchanges ($M=3.2$, $SD=2.52$), $t(1683)=-8.11$, $p<.001$. Additionally, support requesting exchanges had a higher mean ratio of anxiety words ($M=0.43$, $SD=0.78$) than support providing exchanges ($M=0.28$, $SD=0.63$), $t(1683)=4.36$, $p<.001$, as well as a higher mean ratio of anger words ($M=0.3$, $SD=0.71$)

than support providing exchanges ($M=0.23$, $SD=0.67$), $t(1683)=2.03$, $p<.05$, and a higher mean ratio of sadness words ($M=0.49$, $SD=0.86$) than support providing exchanges ($M=0.4$, $SD=0.97$), $t(1683)=2.10$, $p<.05$.

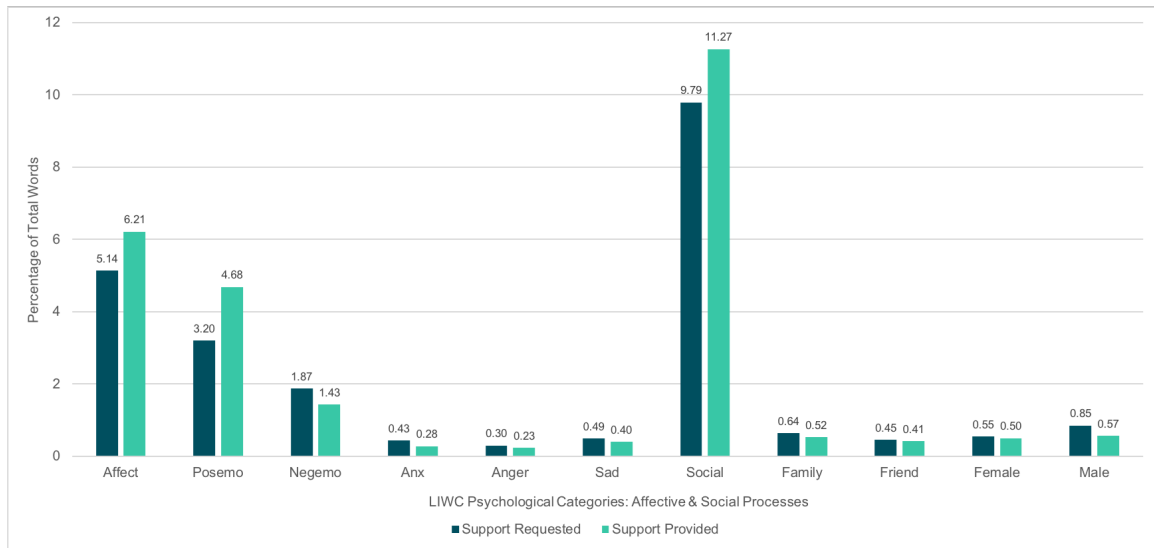


Figure 12. Mean percentage of total words for LIWC psychological categories of affective and social processes by exchanges that requested and provided social support.

Concerning social processes, support requesting exchanges had a higher mean ratio of family words ($M=0.64$, $SD=1.13$) than support providing exchanges ($M=0.52$, $SD=0.99$), $t(1683)=2.16$, $p<.05$, as well as a higher mean ratio of male references words ($M=0.85$, $SD=1.89$) than support providing exchanges ($M=0.57$, $SD=1.32$), $t(1683)=3.64$, $p<.001$. In addition, there were not significant differences between support requesting exchanges and support providing exchanges regarding words reflecting the social processes of friends and female references. Next, Figure 13 shows the mean percentage of total words for LIWC psychological categories of cognitive processes and drives by

exchanges that requested and provided social support. AYA cancer survivors used a higher number of words reflecting cognitive processes, particularly insight, tentative, and differentiation, as well as drives, particularly affiliation.

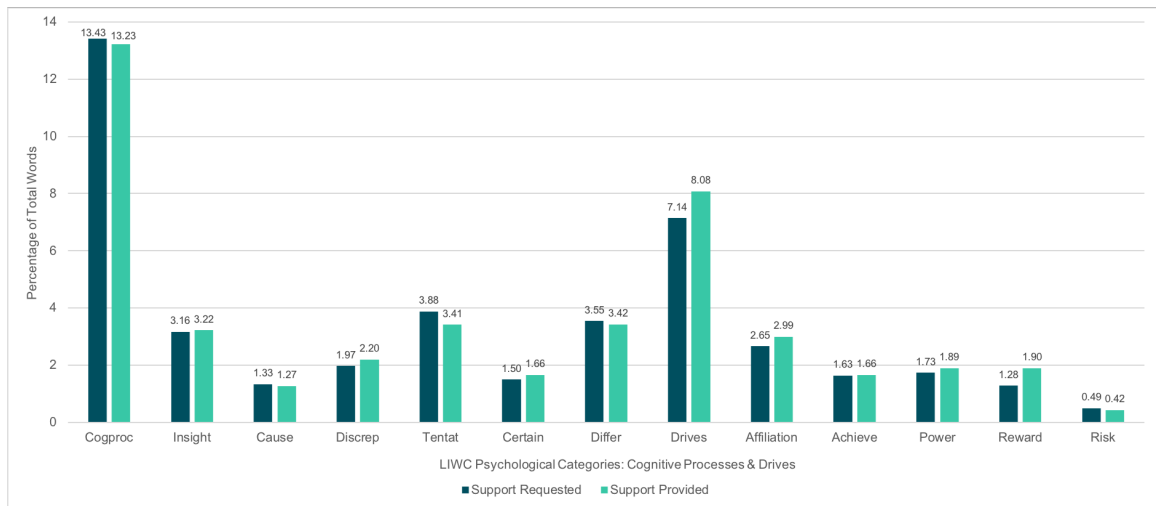


Figure 13. Mean percentage of total words for LIWC psychological categories of cognitive processes and drives by exchanges that requested and provided social support.

As for cognitive processes, support requesting exchanges had a higher mean ratio of tentative words ($M=3.88$, $SD=2.87$) than support providing exchanges ($M=3.41$, $SD=2.58$), $t(1683)=3.41$, $p<.001$, while support providing exchanges had a higher mean ratio of discrepancy words ($M=2.2$, $SD=2.22$) than support requesting exchanges ($M=1.97$, $SD=1.77$), $t(1683)=-2.14$, $p<.05$. Moreover, there were not significant differences between support requesting exchanges and support providing exchanges regarding words reflecting the cognitive processes of insight, causation, certainty, and differentiation.

As for drives, support providing exchanges had a higher mean ratio of affiliation words ($M=2.99$, $SD=3.21$) than support requesting exchanges ($M=2.65$, $SD=2.93$), $t(1683)=-2.15$, $p<.05$, as well as a higher mean ratio of reward words ($M=1.9$, $SD=2.54$) than support requesting exchanges ($M=1.28$, $SD=1.47$), $t(1683)=-5.51$, $p<.001$. In addition, there were no significant differences between support requesting exchanges and support providing exchanges regarding words reflecting drives of achievement, power, and risk.

Last, additional significant differences were found for LIWC linguistic variables related to psychological processes that are worth noting, including time orientations, relativity, personal concerns, and informal language. First, AYA cancer survivors used a higher number of words reflecting time orientations of present and past focus. Support requesting exchanges had a higher mean ratio of words reflecting past focus ($M=4.73$, $SD=3.08$) than support providing exchanges ($M=4.15$, $SD=3.27$), $t(1683)=3.60$, $p<.001$, while support providing exchanges had a higher mean ratio of words reflecting future focus ($M=1.69$, $SD=1.98$) than support requesting exchanges ($M=1.35$, $SD=1.55$), $t(1683)=-3.76$, $p<.001$. In addition, even though the ratio of words reflecting present focus was the highest, no significant differences between support requesting exchanges and support providing exchanges were found. Second, AYA cancer survivors used a higher number of words reflecting relativity, particularly time. Support requesting exchanges had a higher mean ratio of words reflecting time ($M=6.98$, $SD=3.66$) than support providing exchanges ($M=5.89$, $SD=3.87$), $t(1683)=5.66$, $p<.001$. Third, AYA cancer survivors used a higher number of words reflecting personal concerns of work,

leisure, and religion. Support providing exchanges had a higher mean ratio of words reflecting work ($M=1.27$, $SD=2.03$) than support requesting exchanges ($M=1.08$, $SD=1.56$), $t(1683)=-2.01$, $p<.05$, as well as a higher mean ratio of words reflecting religion ($M=0.52$, $SD=3.97$) than support requesting exchanges ($M=0.18$, $SD=3.18$), $t(1683)=-2.09$, $p<.05$. There were no significant differences between support requesting exchanges and support providing exchanges regarding words reflecting personal concerns of leisure.

Finally, AYA cancer survivors used a higher number of words reflecting informal language. Support providing exchanges had a higher mean ratio of words reflecting informal language ($M=1.58$, $SD=2.8$) than support requesting exchanges ($M=0.9$, $SD=1.58$), $t(1683)=-5.51$, $p<.001$. Likewise, support providing exchanges had significantly higher mean ratios of words reflecting netspeak, assent, and nonfluencies than support requesting exchanges. Corresponding mean word ratios as well as t test and p values for these linguistic variables are provided in Appendix E.

4.3.3 Conclusions

The linguistic analysis of supportive messages exchanged by participants of the CSN discussion board “Young Cancer Survivors” indicate that support requesting exchanges contained more words on average, used more words longer than six letters, and used more first-person and third-person singular pronouns. In addition, support requesting exchanges used more negative emotion words, particularly words reflecting anxiety, anger, and sadness. Moreover, support requesting exchanges used more words

related to family topics, male references, tentativeness, and had more words reflecting past focus as well as time.

As for support providing exchanges, participants used more first-person plural pronouns, second-person pronouns, and third-person plural pronouns. In addition, support providing exchanges used more positive emotion words, discrepancy words, words reflecting affiliation, reward, and future focus. Last, support providing exchanges used more words related to topics of work and religion and used more informal language.

**Chapter 5. Specific Aim 2: To Examine Network Structures of Online Interactions
Among AYA Cancer Survivors Based on Content Characteristics Determined in
Specific Aim 1**

The previous chapter provided a characterization of online exchanges among AYA cancer survivors according to social support, SCP domains, psychosocial need themes, and language features. Once these theoretical and empirical concepts have been appropriately documented, it was possible to study the social structure defined by naturally occurring online exchanges. Using a social network analysis approach allowed me to gain an in-depth understanding of the social structures of online supportive interactions among AYA cancer survivors as well as how they influence the exchange of social support in an online context. Hence, SA 2 combines content and linguistic analyses with social network analysis. This analytical approach focused on revealing and assessing the structure of support exchanges among AYA cancer survivors to examine network properties. In addition, a social network analysis approach provides insight into how the properties of network structures, such as the density, size, or number of contacts are associated with different types of social support and linguistic features. Therefore, the associations between support networks as well as their structural properties and linguistic features were assessed. Lastly, insights were gathered from support networks and their associations with structural properties and language features. These insights can inform the design of potentially impactful online support interventions targeting AYA cancer survivors. To achieve SA 2, two objectives were completed and are presented next.

5.1 Objective 1: Characterize Social Network Structures Based on Types of Social Support, SCP Domains, and Psychosocial Need Themes

The characterization of the content of naturally occurring online exchanges among AYA cancer survivors, achieved in SA 1, provided a measure of the types and extent of social support, SCP domains, psychosocial need themes, and linguistic features. In addition to the functional aspects that online support exchanges may have, from a social support perspective, the social network created by AYA cancer survivors while connecting and interacting online is key (Stansfeld, 2006; Holt-Lunstad & Uchino, 2015). For instance, social networks have been used to measure the sources and flows of social support (Marmot & Wilkinson, 2006). In addition, when the functional aspects of support exchanges are known, finer analyses may be achieved by combining both functional aspects and the social network (Heaney & Israel, 2008; Hether et al., 2016; Stansfeld, 2006). Consequently, SA 2 focuses on revealing the social networks defined by the support exchanges among AYA cancer survivors in the CSN discussion board “Young Cancer Survivors”. The structural properties of these networks were examined and assessed based on types of social support, SCP domains, psychosocial need themes, and linguistic features. In this section, I describe such support networks and their structural properties.

5.1.1 Methods

Social network analysis is increasingly used in many scientific disciplines, including biomedical informatics. Social network analysis refers to the integration of network methods into the study of constructs used by theories in each field of study. The

main characteristic of social network analysis is its focus on studying the relationships between individuals instead of the individuals' themselves and their attributes (Valente, 2015). In the context of social support, social networks have been used to measure the sources and flows of social support. For example, social networks can be used to measure the number of individuals in the network, the number of connections between them, or the frequency of contact (Marmot & Wilkinson, 2006). In addition, a structural or social network perspective focuses on "how an individual's social connections provide social support and predict health-related outcomes" (Hether et al., 2016). Consequently, a social network perspective provides insights into the associations between network characteristics, such as density and reciprocity, and social support (Hether et al., 2016). Researchers have suggested that combining the functional content of relationships, i.e., the quality and type of support, with the social networks can provide in-depth analyses (Heaney & Israel, 2008; Hether et al., 2016; Stansfeld, 2006). Therefore, when the functional aspects of support exchanges are known, finer analyses may be achieved by combining both functional aspects (content) and the social networks.

5.1.1.1 Unit of Analysis. The unit of analysis for the social network analysis comprises dyadic (relational) data from online interactions, which includes all the connections between participants of the CSN discussion board "Young Cancer Survivors" as well as the support messages exchanged (support provided and requested). A connection between two participants is defined by an online interaction when a support message was exchanged between them.

5.1.1.2 Data Extraction. Network data was extracted from the CSN discussion board “Young Cancer Survivors” as well as the content analysis as described in SA 1. Network data included: (1) participants’ usernames (anonymized using random numeric identifiers), including both source and target of a message; (2) content analysis variables; (3) the frequency or number of supportive messages exchanged between participants; and (4) the linguistic features.

5.1.1.3 Network Measures. Table 5 contains the network measures used to describe and assess network structures. Gephi (version 0.9.2) and UCINET (version 6.749) were used to create visualizations of the network structures and to calculate network measures (Bastian M. et al., 2009; Borgatti, S.P. et al., 2002).

Table 5

Network structure measures

Network measure	Definition
Size	Number of participants in the network.
Density	Indicates the proportion of connections relative to the total number possible.
Diameter	The length of the longest path connecting two participants in the network.
Average degree	The mean number of connections to and from a participant.
Average in-degree	The mean number of support messages received by a participant.

Network measure	Definition
Average out-degree	The mean number of support messages sent by a participant.
In-degree centralization	Variance of in-degree scores. *
Out-degree centralization	Variance of out-degree scores. **
Average path length	Indicates the mean distances between all participants in the network.
Betweenness centrality	Indicates how often a participant appears on shortest paths between participants in the network.
Closeness centrality	The mean distance from a participant to all other participants in the network.
Eccentricity	The distance from a given participant to the farthest participant from it in the network.
Modularity	Indicates how well a network decomposes into modular communities.
Average clustering coefficient	Mean of the measure of how complete the neighborhood of a participant is (the neighborhood of a participant is the set of participants that are connected to that participant).

* A higher in-degree metric reflects a network that is highly organized around one central individual receiving the most support (Freeman, 1978; Hether et al., 2016).

** A higher out-degree metric reflects a network that is highly organized around one central individual providing the most support (Freeman, 1978; Hether et al., 2016).

5.1.2 Results

Six social networks were created for the CSN discussion board “Young Cancer Survivors”: one multiplex support network depicting overall support exchanges and five uniplex support networks depicting each dimension of support exchanges among AYA cancer survivors. Table 6 shows the descriptive network measures for each support network. The multiplex support network is naturally the largest network consisting of 625 participants with 1021 connections between them. In addition, the multiplex support network is the most active network with an average degree of 1.634 which indicates that each participant has on average less than two connections with other participants. Moreover, the multiplex support network has a low density of .003 indicating that less than 1% of all possible connections in the network were realized. Furthermore, the multiplex support network is slightly more centralized on in-degree (5.84%) suggesting some participants received more messages than others. Figure 14 shows the multiplex support network.



Figure 14. Communication support network.

As for the five uniplex support networks, the emotional support network was the largest network consisting of 491 participants with 713 connections between them, followed by the informational and esteem support networks, consisting of 339 and 310 participants with 414 and 397 connections between them, respectively, and the network and tangible aid support networks, consisting of 132 and 33 participants with 143 and 36 connections between them, respectively. Figures 15-18 show the uniplex support

networks. In addition, the emotional support network is the most active uniplex network with an average degree of 1.452, followed by the esteem and informational support networks, with similar average degrees of 1.281 and 1.221, respectively, and the tangible aid and network support networks, again with similar average degrees of 1.091 and 1.083, respectively. However, despite the emotional network being the most active uniplex network, the average degrees indicate that across all uniplex networks each participant has on average less than two connections with other participants. Also, all uniplex networks have low densities ($<.008$) with the exception of the tangible aid network that has a density of .034, yet indicating that only less than 4% of all possible connections in the network were realized. Moreover, the two largest uniplex networks, emotional and informational support, have weak and very similar in- and out-degree centralization indexes, suggesting that these networks are not strongly centralized on either of these measures. However, the tangible aid, network, and esteem support networks are more centralized on in-degree, indicating that some participants received more tangible aid, network, and esteem support support than others.

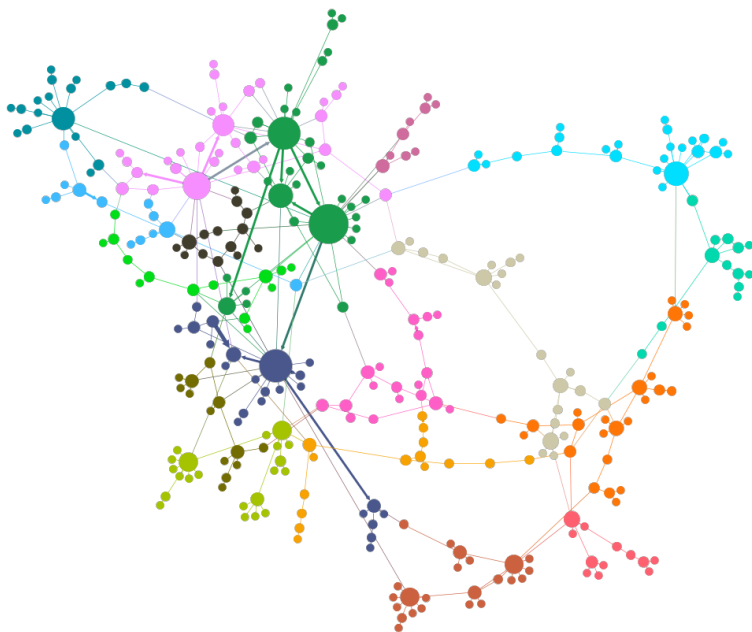


Figure 15. Informational support network.



Figure 16. Emotional support network.

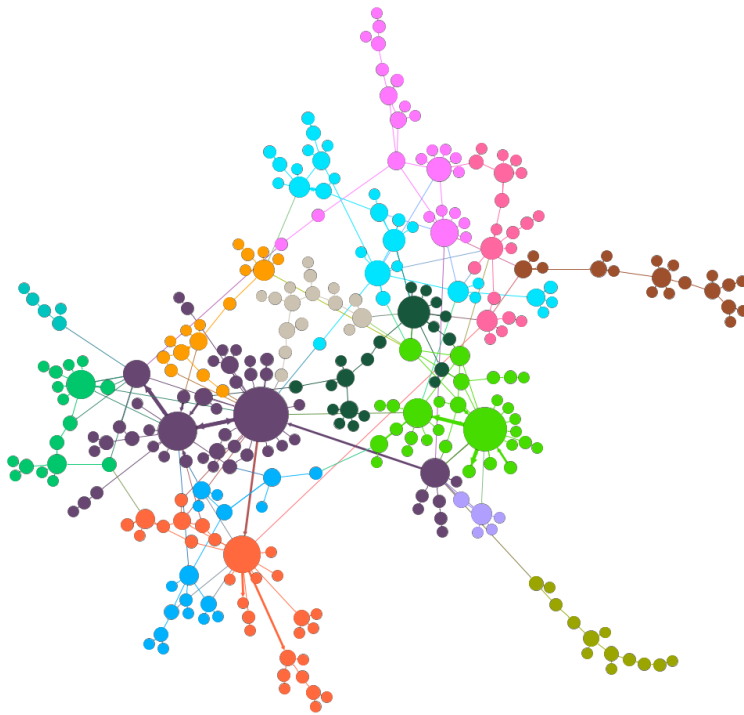


Figure 17. Esteem support network.

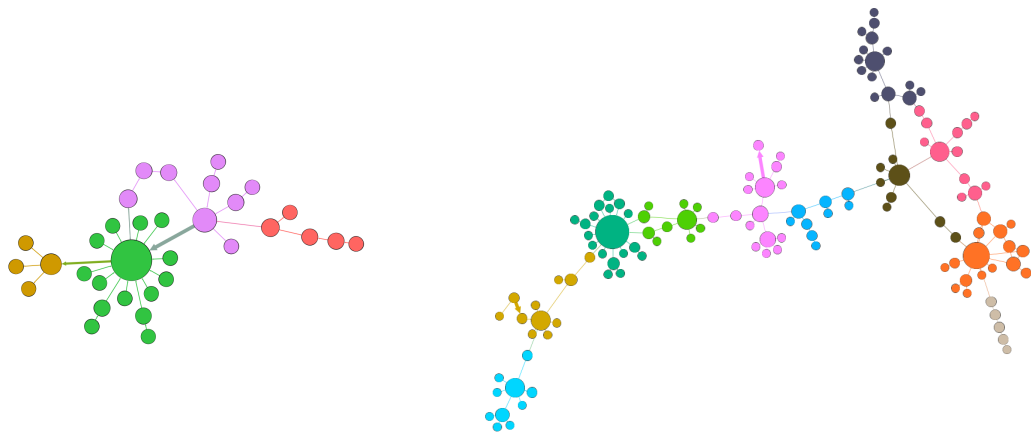


Figure 18. Tangible aid (left) and network (right) support networks.

Table 6

Summary of network metrics for CSN Young Adult forum's social support networks

Network	Multiplex Support	Informational Support	Emotional Support	Esteem Support	Tangible Aid	Network Support
Size	625	339	491	310	33	132
Density	0.003	0.004	0.003	0.004	0.034	0.008
Diameter	14	9	12	15	4	6
Average degree	1.634	1.221	1.452	1.281	1.091	1.083
In-degree centralization	0.0584	0.035	0.038	0.0609	0.4627	0.1078
Out-degree centralization	0.0311	0.032	0.0318	0.0316	0.0967	0.0458
Average path length	4.892	3.922	4.42	4.395	2.495	2.008
Modularity	0.683	0.778	0.742	0.763	0.563	0.796
Average clustering coefficient	0.046	0.023	0.026	0.021	0	0.011

Four two-mode networks were created for the CSN discussion board “Young Cancer Survivors” each one formed by participants and psychosocial need themes representing support exchanges by SCP domains. Figure 19-22 show the two-mode network for the SCP domains (cancer surveillance and screening, late effects/side effects management, risk reduction and cancer prevention, and psychosocial functioning). According to degree centrality, emotions/mental health and social support were the most prevalent psychosocial need themes embedded in the support exchanges among participants for the SCP domain 1 cancer surveillance and screening. In addition, physical, emotions/mental

health, fertility and reproductive health, social support, education/information, and dating and relationships were the most prevalent psychosocial need themes embedded in the support exchanges among participants for the SCP domain 2 late effects/side effects management. Moreover, resources and education/information were the most prevalent psychosocial need themes embedded in the support exchanges among participants for the SCP domain 3 risk reduction and cancer prevention. Lastly, social support, peer support, emotions/mental health, and dating and relationships were the most prevalent psychosocial need themes embedded in the support exchanges among participants for the SCP domain 4 psychosocial functioning.

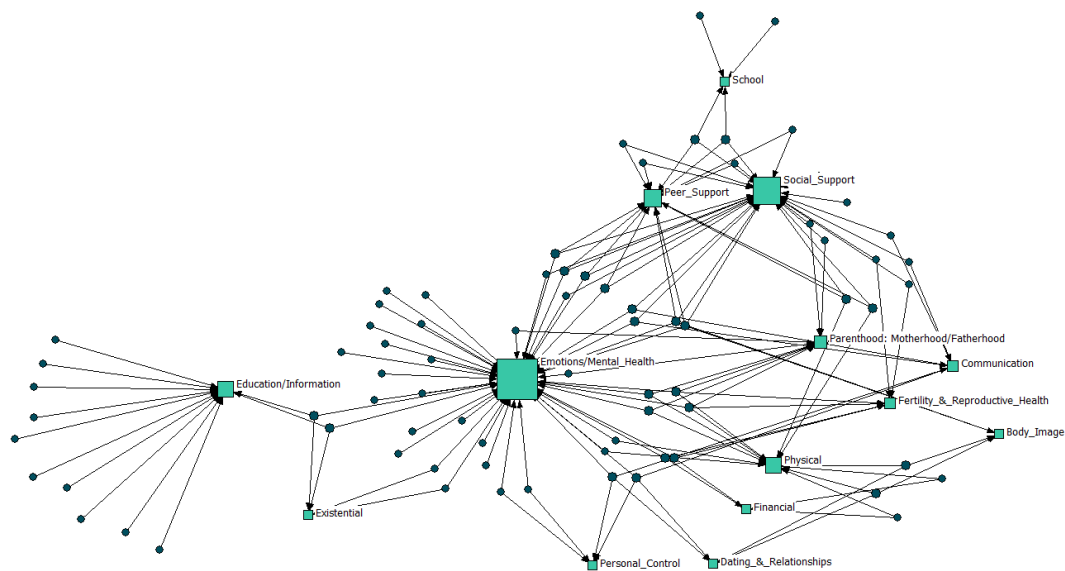


Figure 19. Two-mode network for the SCP domain 1 cancer surveillance and screening, formed by CSN discussion board “Young Cancer Survivors” participants and psychosocial need themes in their support exchanges.

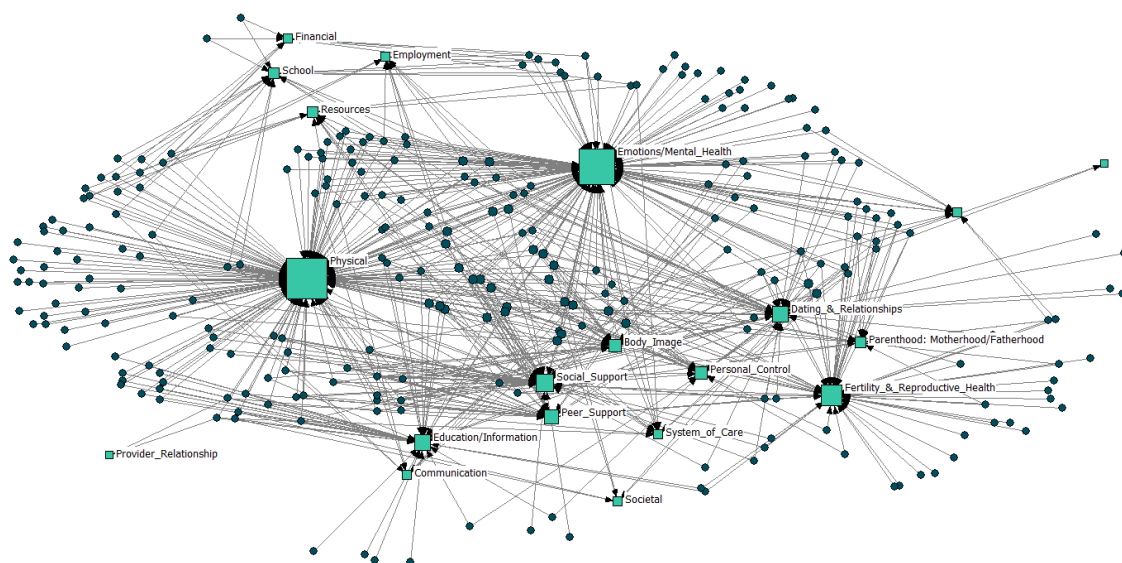


Figure 20. Two-mode network for the SCP domain 2 late effects/side effects management, formed by CSN discussion board “Young Cancer Survivors” participants and psychosocial need themes in their support exchanges.

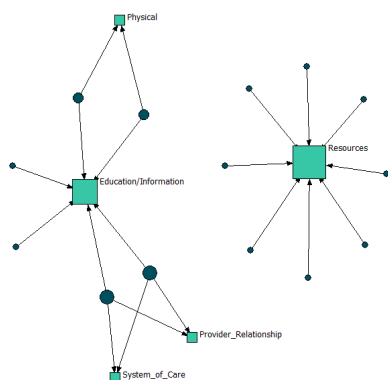


Figure 21. Two-mode network for the SCP domain 3 risk reduction and cancer prevention, formed by CSN discussion board “Young Cancer Survivors” participants and psychosocial need themes in their support exchanges.

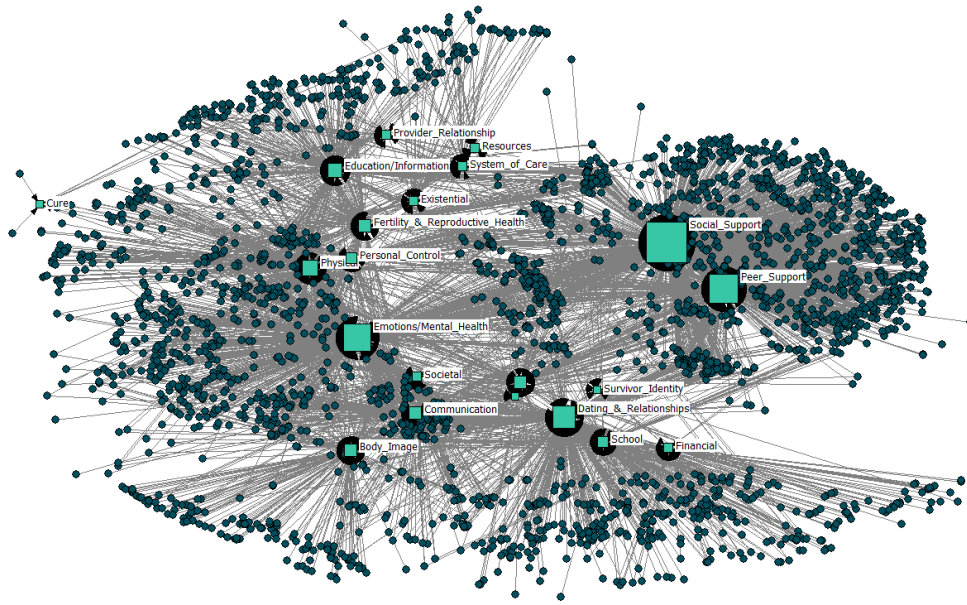


Figure 22. Two-mode network for the SCP domain 4 psychosocial functioning, formed by CSN discussion board “Young Cancer Survivors” participants and psychosocial need themes in their support exchanges.

5.1.3 Conclusions

A social network perspective provided an in-depth understanding of support exchanges among AYA cancer survivors networks. This approach allowed me to examine the structural properties of support networks. All support networks varied in size and were characterized by low densities and low average degrees. In addition, the network structure and community development of support networks show that subcommunities were created as indicated by modularity values. However, average clustering coefficients were low for all support networks, indicating that few support

exchanges took place between participants and that all support networks were comprised of a high number of weak ties.

5.2 Objective 2: Assess the Relationships Between the Structure of Networks Based on Types of Social Support and Linguistic Features

Once the networks formed by the online exchanges of social support among AYA cancer survivors have been revealed and their structural properties assessed, associations between support networks as well as their network properties and linguistic features were assessed. In this section, I describe the methods used to assess network associations as well as the results.

5.2.1 Methods

5.2.1.1 Network Correlations. Researchers have developed methods to statistically test network hypothesis. In addition, network statistical methods have been designed to deal with the important feature of social networks of interdependencies (Borgatti et al., 2018). A common network statistical approach is the permutation test. The permutation test is a nonparametric approach that does not depend on the assumption of data being drawn from a probability distribution and generates statistical distributions from the network data itself, thus overcoming the independence assumption in conventional statistics. Moreover, the permutation test allows testing for associations between networks with the same actors using correlation. Hence, a correlation between two networks formed by different types of relations among the same actors assesses the probability of a tie of one type of relation being related with the probability of another

(Borgatti et al., 2018). Such associations between networks can be calculated with the Quadratic Assignment Procedure (QAP) test (Krackhardt, 1988).

QAP is a nonparametric, permutation-based technique that “correlates the two matrices by effectively reshaping them into two long columns as described above and calculating an ordinary measure of statistical association such as Pearson’s r ” (Borgatti et al., 2018). In addition, the significance of the observed correlation is calculated by comparing “the observed correlation to a reference set of thousands of correlations between thousands of pairs of matrices that are just like the data matrices but are known to be independent of each other” (Borgatti et al., 2018). Subsequently, a p -value is estimated by counting “the proportion of these correlations among independent matrices that were as large as the observed correlation” (Borgatti et al., 2018). Significance is also considered with p -values of less than 5% ($<.05$).

The QAP technique was used to assess associations between support networks. Support networks were assessed on the basis that AYA cancer survivors participating in the CSN discussion board “Young Cancer Survivors” who exchanged a given type of support were likely to exchange another type of support. Accordingly, the following hypothesis was tested:

- $H2$: Dyads of AYA cancer survivors exchanging a given type of support are likely to exchange another type of support.

Thus, support networks were correlated with each other to assess the associations between them. UCINET (version 6.749) was used to perform QAP analyses (Bastian M. et al., 2009; Borgatti, S.P. et al., 2002).

5.2.1.2 Associations Between Support Networks and Linguistic Features.

The linguistic features extracted from the content of online support interactions among AYA cancer survivors, described in Chapter 4, showed their patterns of language use as well as significant differences between support provided and requested. These patterns and differences in language use of AYA cancer survivors are key to gain a deep understanding of their psychological states when they exchange social support online. However, recent studies suggest network structure is also key for the exchange of social support in OSCs. For example, Chen and colleagues (2020) showed that network structural positions occupied by participants of the CSN discussion board “Breast Cancer” were associated with language use among cancer survivors, such as use of words reflecting positive emotion (Chen et al., 2021). In addition, network structural positions occupied by participants of OSCs have been shown to be associated with linguistic features as well as different types of social support (Meng et al., 2016; Xu & Zhang, 2016). Accordingly, results from the linguistic analysis and social network analysis are combined to assess the associations between linguistic features and structural network properties of support networks. The linguistic features included a set of 35 LIWC variables that characterized the language use among AYA cancer survivors to exchange social support online and also showed significant differences between online exchanges that requested social support and online exchanges that provided social support. The structural network properties included: degree, in-degree, out-degree, eccentricity, closeness, betweenness, and clustering coefficient. Pearson correlations were calculated based on online support exchanges and linguistic features for both support providing

(n=475) and requesting (n=343) exchanges among AYA cancer survivors. R (version 4.0.3) was used to calculate Pearson correlations between linguistic features and structural network properties for exchanges that provided and requested support (R Core Team, 2020).

5.2.2 Results

5.2.2.1 Network Correlations. Table 7 shows network correlation scores obtained by performing the QAP analysis. The QAP correlation scores show significant moderate positive correlation between the informational and emotional ($r=.524$; $p<0.000$) and emotional and esteem support networks ($r=.594$; $p<0.000$). Additionally, the informational and esteem ($r=.385$; $p<0.000$) and emotional and network support ($r=.362$; $p<0.000$) networks were significantly positively correlated, although the correlation scores show a much weaker relationship compared with the correlations between the informational and emotional and emotional and esteem support networks. Despite significant weak positive correlations existing between the remaining support networks, correlation scores are small, suggesting that no substantial similarities exist between them.

Table 7

QAP correlation results

	Informational	Emotional	Esteem	Tangible aid	Network
Informational	1	0.524	0.385	0.130	0.204
Emotional	0.524	1	0.594	0.243	0.362

	Informational	Emotional	Esteem	Tangible aid	Network
Esteem	0.385	0.594	1	0.228	0.201
Tangible aid	0.130	0.243	0.228	1	0.158
Network	0.204	0.362	0.201	0.158	1

Note. All correlation scores with $p < 0.000$

5.2.2.2 Associations Between Support Networks and Linguistic Features. In

terms of support providing exchanges among AYA cancer survivors, statistically significant correlations were found between 12 linguistic features and 5 structural network properties. Figure 23 shows a correlation matrix for support providing exchanges depicting the correlations between linguistic features and network structural properties. The color of the cells indicates the Pearson correlation coefficient according the blue-red scale and the asterisks indicate the level of significance (* = $p < .05$; ** = $p < .01$; *** = $p < .001$).

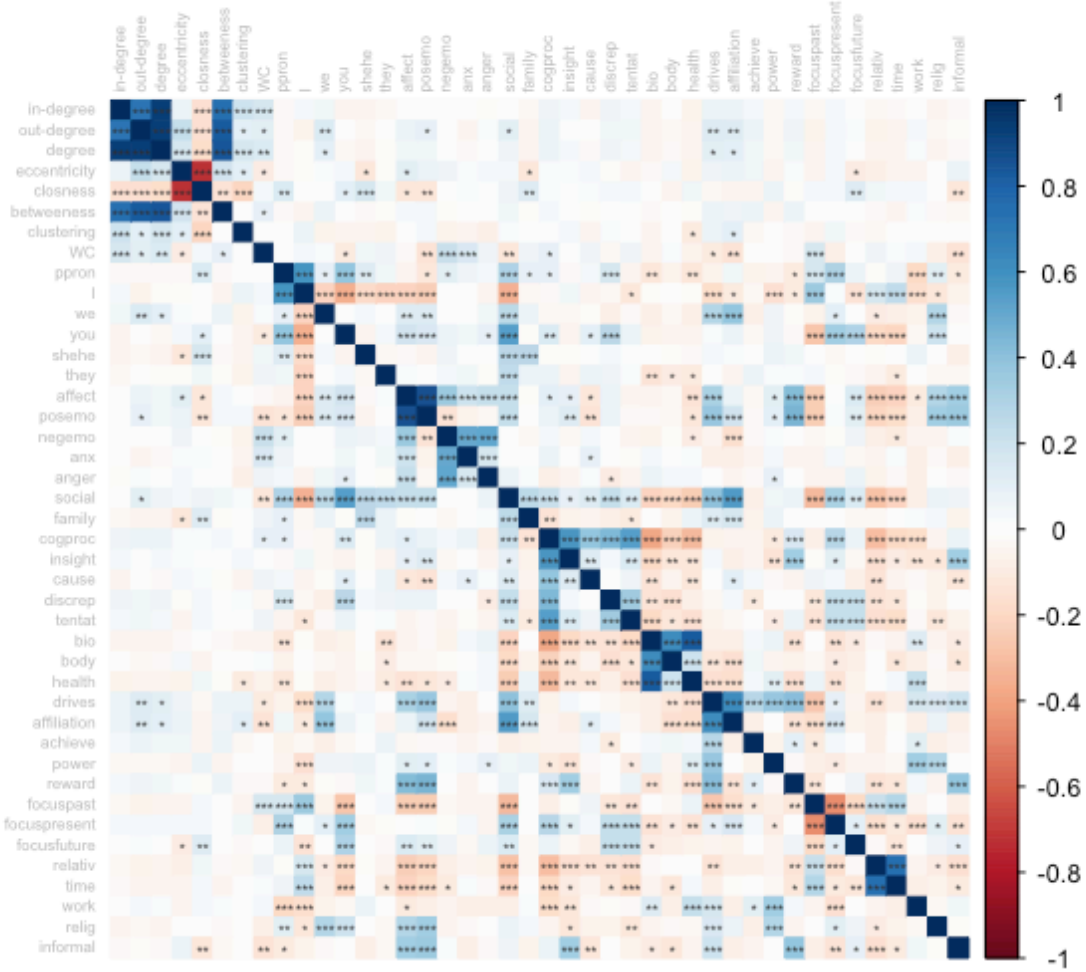


Figure 23. Correlation matrix for support providing exchanges showing the correlations between linguistic features and network structural properties.

In terms of support requesting exchanges among AYA cancer survivors, statistically significant correlations were found between 11 linguistic features and 7 structural network properties. Figure 24 shows a correlation matrix for support requesting messages depicting the correlations between linguistic features and network structural properties.

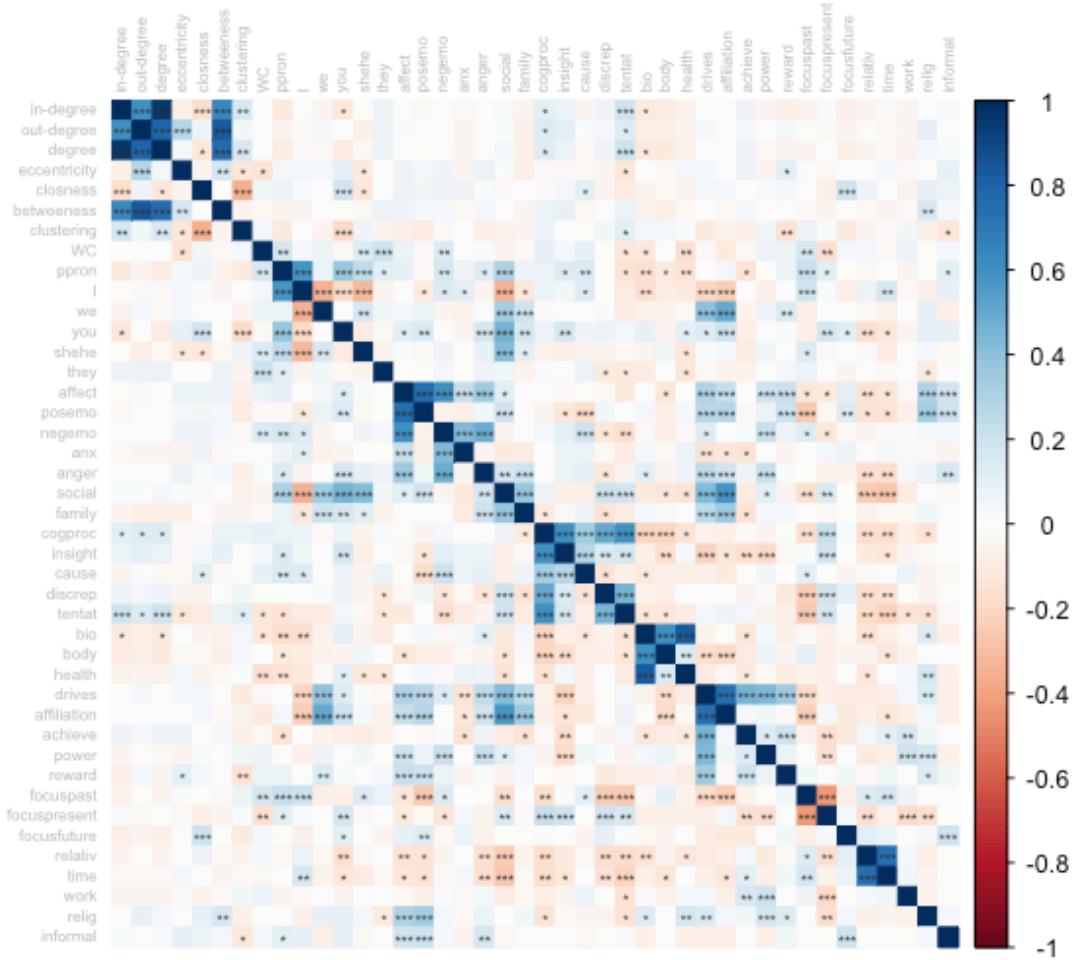


Figure 24. Correlation matrix for support requesting exchanges showing the correlations between linguistic features and network structural properties.

5.2.3 Conclusions

Network correlations showed that the informational and emotional support networks as well as the emotional and esteem support networks are associated. This suggests that AYA cancer survivors that exchange informational and esteem support are likely to exchange emotional support. In addition, correlations show weak associations between support exchanges and linguistic features and structural network properties.

Chapter 6: Discussion and Conclusions

6.1 Discussion for Specific Aim 1: Characterization of the Content of Online Interactions Among AYA Cancer Survivors in Terms of Types of Social Support, Survivorship Care Plan Domains, Psychosocial Needs, and Linguistic Features

The scientifically grounded process followed in Specific Aim 1 resulted in a concise and valid coding scheme that extends previous studies of online support communities targeting cancer survivors. It included Survivorship Care Plan domains as well as a comprehensive taxonomy of psychosocial need themes. SCPs are considered key tools for the delivery of quality survivorship care and were recommended by the former IOM (IOM, 2006; Klemp, 2015; Nekhlyudov et al., 2017; Palos & Gilmore, 2019). Moreover, by incorporating SCP domains, this study extends the study of social support and psychosocial needs among cancer survivors in line with the core elements of cancer survivorship care. As described in Chapter 4, the content analysis results showed that emotional, informational, and esteem support were the most exchanged online among AYA cancer survivors. Despite exchanging more emotional support, AYA cancer survivors exchanged informational and esteem support in similar proportions. These findings are in line with previous studies of online support communities in which informational and emotional support have been found most frequently (Wang et al., 2012; Biyani et al., 2014; Wang et al., 2015; Wang et al., 2015; Wang et al., 2015; Wang et al., 2017). However, these results differ from previous studies in two ways. On the one hand, given that emotional support was considerably more exchanged than informational support, it suggests a high need for emotional support among AYA cancer survivors. On

the other hand, given that esteem support was exchanged in similar proportions with informational support, it suggests that AYA cancer survivors are not only in need of exchanging informational support but also exchanging esteem support to either request or provide compliments, validation, relief of blame and anchorage (i.e. sharing experiences with the community). Furthermore, the results suggest that the amount of social support provided exceeded the number of requests for support among AYA cancer survivors, indicating that online support communities such as the CSN support community are reliable sources of social support, particularly among AYA cancer survivors.

The results from Specific Aim 1 also suggest that most online support exchanges among AYA cancer survivors align with SCP domains. Particularly, the majority of online support exchanges among AYA cancer survivors were aligned with the SCP domain of psychosocial functioning, highlighting a high demand for psychosocial support among members of this population. Given that SCPs are a key communication tool to help personalize and coordinate care between cancer survivors and healthcare providers, this study provides novel and valuable data-driven insights to guide the use of this tool to provide age-specific and developmentally appropriate information and knowledge to help AYA cancer survivors access the support they need and address their unique psychosocial needs (Hydeman et al., 2019; Janssen et al., 2021; Klemp, 2015; Nekhlyudov et al., 2017). In addition, the results show that the number of online support exchanges providing support aligned with the SCP domain of psychosocial functioning slightly exceeded the number of online exchanges that requested support, suggesting that

online support communities can likewise be reliable sources of psychosocial support, particularly for AYA cancer survivors.

As for psychosocial needs, this study shows that a variety of psychosocial needs were discussed online among AYA cancer survivors. Even though most online support exchanges among AYA cancer survivors were related to social support, peer support, emotions and mental health, dating and relationships, physical, and education and information needs, all psychosocial need themes were coded in the dataset. For example, informational support exchanges were related to dating and relationships and education/information, among others; esteem support exchanges were related to emotions/mental health and dating and relationships, among others; and network support exchanges were related to social support and peer support, among others. These findings are consistent with previous research focusing on identifying psychosocial needs for AYA cancer survivors (Choi et al., 2022; Galán et al., 2017, 2018; Lea et al., 2020). However, this study is among the first to use online support communities to identify the psychosocial needs among AYA cancer survivors in the context of online social support exchanging, thus not only highlighting their psychosocial needs but also the types of social support exchanged to address them. Additionally, the results show that psychosocial need themes likewise align with SCP domains. For example, psychosocial need themes aligned with the SCP domain of late effects/side effects management include physical and emotions/mental health. These data-driven insights are key to guide the use of SCPs to improve the provision of relevant information as well as the communication between healthcare providers and AYA cancer survivors regarding their unique

psychosocial needs to help them transition back to normal life (Hydeman et al., 2019; Janssen et al., 2021; Klemp, 2015; Nekhlyudov et al., 2017).

Finally, the results from the computerized text analysis provide valuable insights into how AYA cancer survivors are using language to exchange social support online. Overall, online support exchanges indicate that AYA cancer survivors use words that reflect authenticity, affective, and social and cognitive processes. These results indicate that participants perceive the online community as a safe space to communicate openly and honestly, and to find emotional support and insight, which is consistent with existing research specifically on AYA cancer survivors (Crook et al., 2016; Kaal et al., 2018; Thompson et al., 2016). However, this study is the first to examine linguistic features between online exchanges that requested and provided social support among AYA cancer survivors. As described in Chapter 4, the analysis revealed statistically significant differences in language use between online exchanges that requested and provided social support. On the one hand, online exchanges that requested social support contained more words on average, used more words longer than six letters, more first-person and third-person singular pronouns, and used more negative emotion words, particularly words reflecting anxiety, anger, and sadness. On the other hand, online exchanges that provided social support used more first-person plural pronouns, second-person pronouns, and third-person plural pronouns, positive emotion words, discrepancy words, words reflecting affiliation, reward, and future focus. Additionally, they used more words about related to topics of work and religion, and used more informal language. Group differences suggest these language features may be normative in the exchange of social

support among AYA cancer survivors in online support communities. These insights may inform the design and implementation of online support interventions that aim to address the unmet psychosocial needs of AYA cancer survivors.

6.2 Discussion for Specific Aim 2: Examination of Network Structures of Online Interactions Among AYA Cancer Survivors Based on Content Characteristics Determined in Specific Aim 1

The Specific Aim 2 of this study used a social network analysis approach that provided an in-depth understanding of online social support exchanges among AYA cancer survivors. First, by creating networks it was possible to examine the structural properties of support networks. All support networks varied in size and were characterized by low densities and low average degrees. On the one hand, low densities show the low activity levels of the discussion board and also indicate that support networks formed by AYA cancer survivors have low cohesion levels, suggesting that participants exchanged support with few participants. On the other hand, low average degrees indicate that participants exchanged support with only one participant on average. In addition, the three smallest networks were more centralized than the larger networks. Specifically, the esteem, tangible, and network support networks were more centralized in in-degree, indicating that some participants received more esteem, tangible, and network support than others. This suggests that online support communities such as the CSN support community may benefit individuals in need of esteem, tangible, and network support.

In terms of network structure and community development, all support networks show that subcommunities were created as indicated by modularity values, suggesting that subcommunities of network participants exchanging all types of social support reached high levels of development. However, average clustering coefficients were low for all support networks, indicating that few support exchanges took place between participants and that all support networks were comprised of a high number of weak ties. While most online support communities are characterized by weak ties, previous research has commonly associate weak ties with the exchange of informational support (Chang, 2009; Granovetter, 1973; Hether et al., 2016). In contrast, it is an interesting finding of this work that AYA cancer survivors formed weak ties support networks to exchange a substantial amount of emotional support as well as similar amounts of esteem support and informational support. Finally, the informational and emotional support networks as well as the emotional and esteem support networks were associated as indicated by the network correlations. This suggests that AYA cancer survivors that exchange informational support are likely to exchange emotional support. Similarly, AYA cancer survivors that exchange emotional support are likely to exchange esteem support. These findings may result from the nature of online communications in the discussion forum where emotional support was found to be embedded within online support exchanges that provided informational and esteem support. Lastly, despite statistically significant correlations found between a set of linguistic features and structural network properties for both support exchanges that provided and requested support, all correlation coefficients were weak and more research is needed to further explore these associations.

6.3 Limitations

This study is not without limitations. First, the dataset that was used represented a relatively low number of postings, given the context of online communities. This may be due to the use of other social media platforms, such as Facebook, or other support communities by AYA cancer survivors to seek psychosocial support. In addition, the dataset makes it impossible to measure the engagement of lurkers accessing social support. Furthermore, demographic data was not available from the data. Lastly, given that this study focused on one online community within the context of the United States, the results are not generalizable to the entire population of online support communities.

6.4 Conclusions

In conclusion, the informatics approach used in this study provided an in-depth understanding of online support exchanges among AYA cancer survivors. The results show that AYA cancer survivors are mostly exchanging emotional support but exchange informational and esteem support in similar proportions. In addition, most support interactions among AYA cancer survivors align with the psychosocial functioning domain of survivorship care plans, highlighting a high demand for psychosocial support among members of this population. In addition, this study expands current understanding of how AYA cancer survivors are using language to exchange support online, highlighting differences in language use between online requests and provisions of social support. Furthermore, a social network analysis approach examined the structural properties of support networks. A social network approach shows the structural characteristics of support networks such as size differences and low densities and average

degrees across them. Moreover, this study shows that subcommunities of network support developed among AYA cancer survivors, in spite of low levels of cohesion and clustering between them. Interestingly, all support networks were comprised of a high number of weak ties. Additionally, this study shows that AYA cancer survivors who exchange informational or esteem support are also likely to exchange emotional support. Lastly, the novel data-driven insights gathered by applying an informatics approach may inform the design and implementation of online support interventions that aim to address the unmet psychosocial needs of AYA cancer survivors.

6.4.1 Innovation

This study applied an informatics approach to examine online support exchanges among AYA cancer survivors. This informatics approach combined content analysis, computerized text analysis, and social network analysis. Although the use of these informatics techniques is not new to the study of online support communities, this work is the first to apply this approach to examine the understudied population of AYA cancer survivors. In addition, the incorporation of Survivorship Care Plan domains is an important innovative aspect. By using SCP domains, this study leveraged the two commonly siloed practice domains of cancer survivorship care and public health. Moreover, it shed light on how AYA cancer survivors use language to exchange support online, particularly how they request and provide social support.

6.4.2 Contributions to Biomedical Informatics

The informatics approach applied combined diverse techniques to analyze data from online social support exchanges of an understudied cancer survivor population. By

applying such informatics approach to examine behavioral trace data from an online support community, novel insights were extracted on how AYA cancer survivors request and provide social support online. These data-driven insights may inform the design and implementation of online support interventions that aim to address the unmet psychosocial needs of AYA cancer survivors, for example the access to age-specific and developmentally appropriate information and knowledge. Therefore, this study represents an effort to bridge the gap between health and public health researchers and their needs for informatics techniques in the study of online support communities, particularly those targeting AYA cancer survivors.

References

- Abendschein, B. (2020). Uncertainty management in online support forums for stroke survivors and caregivers. In *Social Support and Health in the Digital Age* (pp. 149–169). Lexington Books.
- Adolescent and Young Adult Oncology Progress Review Group. (2006). *Closing the Gap: Research and Care Imperatives for Adolescents and Young Adults with Cancer*. https://www.livestrong.org/sites/default/files/what-we-do/reports/ayao_prg_report_2006_final.pdf
- Altman, D. G. (1991). *Practical statistics for medical research*. London; New York. Chapman and Hall.
- Barnett, M., McDonnell, G., DeRosa, A., Schuler, T., Philip, E., Peterson, L., Touza, K., Jhanwar, S., Atkinson, T. M., & Ford, J. S. (2016). Psychosocial outcomes and interventions among cancer survivors diagnosed during adolescence and young adulthood (AYA): A systematic review. *Journal of Cancer Survivorship: Research and Practice*, 10(5), 814–831. <https://doi.org/10.1007/s11764-016-0527-6>
- Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2018). *Analyzing Social Networks* (Second edition). Los Angeles. SAGE Publications Ltd.
- Boyd, R. L., & Schwartz, H. A. (2021). Natural Language Analysis and the Psychology of Verbal Behavior: The Past, Present, and Future States of the Field. *Journal of Language and Social Psychology*, 40(1), 21–41. <https://doi.org/10.1177/0261927X20967028>

- Centers for Disease Control and Prevention. (2021). *U.S. Cancer Statistics: Highlights from 2018 Incidence* (USCS Data Brief No. 23). <https://www.cdc.gov/cancer/uscs/about/data-briefs/no23-USCS-highlights-2018-incidence.htm>
- Chang, H.-J. (2009). Online Supportive Interactions: Using a Network Approach to Examine Communication Patterns Within a Psychosis Social Support Group in Taiwan. *Journal of the American Society for Information Science and Technology*, 60(7), 1504–1517. Scopus. <https://doi.org/10.1002/asi.21070>
- Chao, C., Bhatia, S., Xu, L., Cannavale, K. L., Wong, F. L., Huang, P.-Y. S., Cooper, R., & Armenian, S. H. (2020). Chronic Comorbidities Among Survivors of Adolescent and Young Adult Cancer. *Journal of Clinical Oncology*, 38(27), 3161–3174. <https://doi.org/10.1200/JCO.20.00722>
- Chen, M., Bell, R. A., & Barnett, G. A. (2021). From Network Positions to Language Use: Understanding the Effects of Brokerage and Closure Structures from a Linguistic Perspective. *Health Communication*, 36(8), 1001–1008. <https://doi.org/10.1080/10410236.2020.1731776>
- Choi, E., Becker, H., & Kim, S. (2022). Unmet needs in adolescents and young adults with cancer: A mixed-method study using social media. *Journal of Pediatric Nursing*, 64, 31–41. <https://doi.org/10.1016/j.pedn.2022.01.019>
- Close, A. G., Dreyzin, A., Miller, K. D., Seynnaeve, B. K. N., & Rapkin, L. B. (2019). Adolescent and young adult oncology—Past, present, and future. *CA: A Cancer Journal for Clinicians*, 69(6), 485–496. <https://doi.org/10.3322/caac.21585>

- Crook, B., Glowacki, E. M., Love, B., Jones, B. L., Macpherson, C. F., & Johnson, R. H. (2016). Hanging by a thread: Exploring the features of nonresponse in an online young adult cancer survivorship support community. *Journal of Cancer Survivorship: Research and Practice*, 10(1), 185–193. <https://doi.org/10.1007/s11764-015-0465-8>
- Egbert, N., & Wright, K. B. (2020). *Social Support and Health in the Digital Age*. Rowman & Littlefield.
- Essue, B. M., Iragorri, N., Fitzgerald, N., & de Oliveira, C. (2020). The psychosocial cost burden of cancer: A systematic literature review. *Psycho-Oncology*, 29(11), 1746–1760. <https://doi.org/10.1002/pon.5516>
- Galán, S., de la Vega, R., & Miró, J. (2018). Needs of adolescents and young adults after cancer treatment: A systematic review. *European Journal of Cancer Care*, 27(6), e12558. <https://doi.org/10.1111/ecc.12558>
- Galán, S., de la Vega, R., Tomé Pires, C., Racine, M., Solé, E., Jensen, M. P., & Miró, J. (2017). What are the needs of adolescents and young adults after a cancer treatment? A Delphi study. *European Journal of Cancer Care*, 26(2). <https://doi.org/10.1111/ecc.12488>
- Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6), 1360–1380.
- Griswold Home Care. (n.d.). *Types of Caregivers*. Retrieved November 12, 2021, from <https://www.griswoldhomecare.com/be-a-caregiver/who-are-caregivers/-types-of-caregivers/>

- Heaney, C. A., & Israel, B. A. (2008). Social Networks and Social Support. In *Health Behavior and Health Education: Theory, Research, and Practice* (4th ed). John Wiley & Sons.
- Hether, H. J., Murphy, S. T., & Valente, T. W. (2016). A social network analysis of supportive interactions on prenatal sites. *Digital Health*, 2.
<https://doi.org/10.1177/2055207616628700>
- HHS (U.S. Department of Health and Human Services). (2012). *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*. Washington, D.C.
<https://www.hhs.gov/guidance/document/guidance-regarding-methods-de-identification-protected-health-information-accordance-0>
- Holland, L. R., Walker, R., Henney, R., Cashion, C. E., & Bradford, N. K. (2021). Adolescents and Young Adults with Cancer: Barriers in Access to Psychosocial Support. *Journal of Adolescent and Young Adult Oncology*, 10(1), 46–55.
<https://doi.org/10.1089/jayao.2020.0027>
- Houlihan, N. G. (2009). Transitioning to cancer survivorship: Plans of care. *Oncology (Williston Park, N.Y.)*, 23(8 Suppl), 42–48.
- Hydeman, J. A., Uwazurike, O. C., Adeyemi, E. I., & Beaupin, L. K. (2019). Survivorship Needs of Adolescent and Young Adult Cancer Survivors: A Concept Mapping Analysis. *Journal of Cancer Survivorship : Research and Practice*, 13(1), 34–42.
<https://doi.org/10.1007/s11764-018-0725-5>

- Institute of Medicine. (2015). *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*. Washington, DC. The National Academies Press. <https://doi.org/10.17226/18998>
- Institute of Medicine (US) Committee on Psychosocial Services to Cancer Patients/Families in a Community Setting. (2008). *Cancer Care for the Whole Patient: Meeting Psychosocial Health Needs* (N. E. Adler & A. E. Page, Eds.). National Academies Press (US).
<http://www.ncbi.nlm.nih.gov/books/NBK4015/>
- Janssen, S. H. M., van der Graaf, W. T. A., van der Meer, D. J., Manten-Horst, E., & Husson, O. (2021). Adolescent and Young Adult (AYA) Cancer Survivorship Practices: An Overview. *Cancers*, 13(19), 4847. <https://doi.org/10.3390/cancers13194847>
- Jin, Z., Griffith, M. A., & Rosenthal, A. C. (2021). Identifying and Meeting the Needs of Adolescents and Young Adults with Cancer. *Current Oncology Reports*, 23(2), 17.
<https://doi.org/10.1007/s11912-020-01011-9>
- Klemp, J. R. (2015). Survivorship care planning: One size does not fit all. *Seminars in Oncology Nursing*, 31(1), 67–72. <https://doi.org/10.1016/j.soncn.2014.11.008>
- Kornfield, R., Toma, C. L., Shah, D. V., Moon, T. J., & Gustafson, D. H. (2018). What Do You Say Before You Relapse? How Language Use in a Peer-to-peer Online Discussion Forum Predicts Risky Drinking among Those in Recovery. *Health Communication*, 33(9), 1184–1193. <https://doi.org/10.1080/10410236.2017.1350906>
- Krackhardt, D. (1988). Predicting with networks: Nonparametric multiple regression analysis of dyadic data. *Social Networks*, 10(4), 359–381. [https://doi.org/10.1016/0378-8733\(88\)90004-4](https://doi.org/10.1016/0378-8733(88)90004-4)

- Lea, S., Martins, A., Fern, L. A., Bassett, M., Cable, M., Doig, G., Morgan, S., Soanes, L., Whelan, M., & Taylor, R. M. (2020). The support and information needs of adolescents and young adults with cancer when active treatment ends. *BMC Cancer*, 20(1), 697. <https://doi.org/10.1186/s12885-020-07197-2>
- Lewallen, A. C., Owen, J. E., Bantum, E. O., & Stanton, A. L. (2014). How language affects peer responsiveness in an online cancer support group: Implications for treatment design and facilitation. *Psycho-Oncology*, 23(7), 766–772. <https://doi.org/10.1002/pon.3477>
- Li, A., Jiao, D., & Zhu, T. (2018). Detecting depression stigma on social media: A linguistic analysis. *Journal of Affective Disorders*, 232, 358–362. <https://doi.org/10.1016/j.jad.2018.02.087>
- Lorgelly, P. K., & Neri, M. (2018). Survivorship burden for individuals, households and society: Estimates and methodology. *Journal of Cancer Policy*, 15, 113–117. <https://doi.org/10.1016/j.jcpo.2018.02.005>
- Meng, J., Chung, M., & Cox, J. (2016). Linking Network Structure to Support Messages: Effects of Brokerage and Closure on Received Social Support. *Journal of Communication*, 66(6), 982–1006. <https://doi.org/10.1111/jcom.12268>
- Miller, K. D., Fidler-Benaoudia, M., Keegan, T. H., Hipp, H. S., Jemal, A., & Siegel, R. L. (2020). Cancer Statistics for Adolescents and Young Adults, 2020. *CA: A Cancer Journal for Clinicians*, 70(6), 443–459. <https://doi.org/10.3322/caac.21637>
- Miller, K. D., Nogueira, L., Mariotto, A. B., Rowland, J. H., Yabroff, K. R., Alfano, C. M., Jemal, A., Kramer, J. L., & Siegel, R. L. (2019). Cancer treatment and survivorship

statistics, 2019. *CA: A Cancer Journal for Clinicians*, 69(5), 363–385.

<https://doi.org/10.3322/caac.21565>

Moessner, M., Feldhege, J., Wolf, M., & Bauer, S. (2018). Analyzing big data in social media:

Text and network analyses of an eating disorder forum. *The International Journal of Eating Disorders*, 51(7), 656–667. <https://doi.org/10.1002/eat.22878>

National Cancer Institute. (n.d.). *Cancer Among Adolescents and Young Adults (AYAs)*—

Cancer Stat Facts. SEER. Retrieved May 9, 2022, from

<https://seer.cancer.gov/statfacts/html/aya.html>

Nekhlyudov, L., Ganz, P. A., Arora, N. K., & Rowland, J. H. (2017). Going Beyond Being

Lost in Transition: A Decade of Progress in Cancer Survivorship. *Journal of Clinical Oncology*, 35(18), 1978–1981. <https://doi.org/10.1200/JCO.2016.72.1373>

O’Dea, B., Larsen, M. E., Batterham, P. J., Calex, A. L., & Christensen, H. (2017). A

Linguistic Analysis of Suicide-Related Twitter Posts. *Crisis*, 38(5), 319–329.

<https://doi.org/10.1027/0227-5910/a000443>

Palos, G. R., & Gilmore, K. R. (2019). Tools and Resources to Improve Cancer Survivorship

Clinical Practice and Care. In M. A. Rodriguez & Lewis E Foxhall (Eds.), *Handbook of Cancer Survivorship Care*. Springer Publishing Company.

[https://connect.springerpub.com/content/book/978-0-8261-3825-](https://connect.springerpub.com/content/book/978-0-8261-3825-5/part/part01/chapter/ch01)

[5/part/part01/chapter/ch01](https://connect.springerpub.com/content/book/978-0-8261-3825-5/part/part01/chapter/ch01)

Park, J., & Look, K. A. (2019). Health Care Expenditure Burden of Cancer Care in the United

States. *Inquiry: A Journal of Medical Care Organization, Provision and Financing*, 56.

<https://doi.org/10.1177/0046958019880696>

- Penn, A., & Kuperberg, A. (2018). Psychosocial Support in Adolescents and Young Adults With Cancer. *Cancer Journal (Sudbury, Mass.)*, 24(6), 321–327.
<https://doi.org/10.1097/PPO.0000000000000339>
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The Development and Psychometric Properties of LIWC2015*.
<https://repositories.lib.utexas.edu/handle/2152/31333>
- Perez, G. K., Salsman, J. M., Fladeboe, K., Kirchhoff, A. C., Park, E. R., & Rosenberg, A. R. (2020). Taboo Topics in Adolescent and Young Adult Oncology: Strategies for Managing Challenging but Important Conversations Central to Adolescent and Young Adult Cancer Survivorship. *American Society of Clinical Oncology Educational Book. American Society of Clinical Oncology. Annual Meeting*, 40, 1–15.
https://doi.org/10.1200/EDBK_279787
- Rodriguez, M. A., & Lewis-Patterson, P. (2019). Defining Cancer Survivorship. In M. A. Rodriguez & Lewis E Foxhall (Eds.), *Handbook of Cancer Survivorship Care*. Springer Publishing Company. <https://connect.springerpub.com/content/book/978-0-8261-3825-5/part/part01/chapter/ch01>
- Shay, L. A., Allicock, M., & Li, A. (2021). “Every day is just kind of weighing my options.” Perspectives of young adult cancer survivors dealing with the uncertainty of the COVID-19 global pandemic. *Journal of Cancer Survivorship: Research and Practice*.
<https://doi.org/10.1007/s11764-021-01069-9>
- Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer statistics, 2022. *CA: A Cancer Journal for Clinicians*, 72(1), 7–33. <https://doi.org/10.3322/caac.21708>

Stansfeld, S. A. (2006). Social support and social cohesion. In *Social Determinants of Health* (2nd ed.). Oxford. Oxford University Press.

<https://doi.org/10.1093/acprof:oso/9780198565895.003.08>

Stepan, K., & Waldmann, L. (2019). Psychosocial Issues. In M. A. Rodriguez & Lewis E Foxhall (Eds.), *Handbook of Cancer Survivorship Care*. Springer Publishing Company.

<https://connect.springerpub.com/content/book/978-0-8261-3825-5/part/part01/chapter/ch01>

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, 29(1), 24–54. <https://doi.org/10.1177/0261927X09351676>

Thompson, C. M., Crook, B., Love, B., Macpherson, C. F., & Johnson, R. (2016).

Understanding how adolescents and young adults with cancer talk about needs in online and face-to-face support groups. *Journal of Health Psychology*, 21(11), 2636–2646.

<https://doi.org/10.1177/1359105315581515>

Valente, T. W. (2015). Social Networks and Health Behavior. In *Health behavior: Theory, research, and practice* (5th ed, pp. 205–222). Jossey-Bass/Wiley.

Warner, E. L., Ellington, L., Kirchhoff, A. C., & Cloyes, K. G. (2018). Acquisition of Social Support and Linguistic Characteristics of Social Media Posts About Young Adult Cancer. *Journal of Adolescent and Young Adult Oncology*, 7(2), 196–203.

<https://doi.org/10.1089/jayao.2017.0076>

- Xu, R., & Zhang, Q. (2016a). Understanding Online Health Groups for Depression: Social Network and Linguistic Perspectives. *Journal of Medical Internet Research*, 18(3), e63. <https://doi.org/10.2196/jmir.5042>
- Xu, R., & Zhang, Q. (2016b). Understanding Online Health Groups for Depression: Social Network and Linguistic Perspectives. *Journal of Medical Internet Research*, 18(3), e63. <https://doi.org/10.2196/jmir.5042>
- Yabroff, K. R., Mariotto, A., Tangka, F., Zhao, J., Islami, F., Sung, H., Sherman, R. L., Henley, S. J., Jemal, A., & Ward, E. M. (2021). Annual Report to the Nation on the Status of Cancer, Part 2: Patient Economic Burden Associated With Cancer Care. *JNCI: Journal of the National Cancer Institute*, 113(12), 1670–1682. <https://doi.org/10.1093/jnci/djab192>
- Zhang, S., O'Carroll Bantum, E., Owen, J., Bakken, S., & Elhadad, N. (2017). Online cancer communities as informatics intervention for social support: Conceptualization, characterization, and impact. *Journal of the American Medical Informatics Association: JAMIA*, 24(2), 451–459. <https://doi.org/10.1093/jamia/ocw093>

Appendix A

Altman's benchmarking guideline for reliability coefficients

Agreement Statistic	Criteria
< 0.20	Poor
0.21 to 0.40	Fair
0.41 to 0.60	Moderate
0.61 to 0.80	Good
0.81 to 1.00	Very Good

Note. Reprinted from "Practical statistics for medical research" by Altman (Altman, 1991).

Appendix B

Inter-rater reliability and level of measurement for study variables

Variable	Level of Measurement	Percent Agreement	Dwet's AC1
Informational Support	Nominal/binomial	74.12	0.56
Emotional Support	Nominal/binomial	77.64	0.67
Esteem Support	Nominal/binomial	74.12	0.58
Tangible Aid	Nominal/binomial	94.11	0.93
Network Support	Nominal/binomial	83.52	0.77
Negative Behaviors	Nominal/binomial	95.88	0.95
Social Support Requested	Nominal/binomial	88.23	0.80
Social Support Provided	Nominal/binomial	68.23	0.36
SCP Domain 1: Cancer surveillance and screening	Nominal/binomial	99.41	0.99
SCP Domain 2: Late effects/side effects management	Nominal/binomial	82.94	0.75
SCP Domain 3: Risk reduction and	Nominal/binomial	94.11	0.94

Variable	Level of Measurement	Percent Agreement	Dwet's AC1
cancer prevention			
SCP Domain 4: Psychosocial	Nominal/binomial	71.76	0.45
functioning			
Physical	Nominal/binomial	81.76	0.74
Financial	Nominal/binomial	98.23	0.98
Education/information	Nominal/binomial	87.05	0.85
Personal control	Nominal/binomial	94.11	0.94
System of care	Nominal/binomial	98.23	0.98
Resources	Nominal/binomial	97.64	0.98
Emotions/mental health	Nominal/binomial	85.88	0.81
Social support	Nominal/binomial	78.23	0.60
Societal	Nominal/binomial	97.05	0.97
Communication	Nominal/binomial	91.76	0.91
Provider relationship	Nominal/binomial	97.64	0.97
Cure	Nominal/binomial	98.23	0.98
Body image	Nominal/binomial	96.47	0.96
Survivor identity	Nominal/binomial	95.29	0.95
Employment	Nominal/binomial	100.0	1.00
Existential	Nominal/binomial	92.94	0.92
Peer support	Nominal/binomial	91.17	0.88

Variable	Level of Measurement	Percent Agreement	Dwet's AC1
Dating and relationship	Nominal/binomial	95.29	0.94
School	Nominal/binomial	99.41	0.99
Cancer survivor	Nominal/binomial	87.64	0.79
Family caregiver	Nominal/binomial	98.82	0.99
Informal caregiver	Nominal/binomial	100.0	1.00
Other: CSN staff/moderator	Nominal/binomial	100.0	1.00
Other: Undiagnosed patient	Nominal/binomial	98.23	0.98
Unknown	Nominal/binomial	87.05	0.82

Appendix C

Code scheme

Social Support			
Support category	Definition	Support subcategory	Purpose of communication
Informational support	Providing information about the problem, how to appraise the problem, or how to cope with the problem.	Suggestion and advice	Offers ideas and suggests actions.
		Referral	Refers the recipient to some other source of help.
		Situation appraisal	Reassesses or redefines the situation.
		Teaching	Teach how to do something or teach facts. Provides detailed information, facts, or news about the situation

or about skills needed to deal with the situation.

Tangible assistance	Offering assistance or tangible resources to help solve the problem. Providing, or offering, to provide goods or services.	Loan	Offer money or material object.
			Offers to lend the recipient something.
		Direct task	Offer to do something related to the problem.
		Indirect task	Offers to take over one or more of the recipient's other responsibilities while the recipient is under stress.
		Active participation	Offers to join the recipient in action that reduces the stress.
		Willingness	Expresses willingness to help any time.
		Complies with request	Agrees to do something after stressed person

requests it.

Esteem support	Communicating confidence in the other's worth, competence, or ability to solve the problem. Communicating respect and confidence in abilities.	Compliment	Says positive things about the recipient or emphasizes the recipient's abilities.
		Validation	Agree with and take other's side. Expresses agreement with the recipient's perspective on the situation.
		Relief of blame	Say it's not other's fault. Tries to alleviate the recipient's feelings of guilt about the situation.
		Anchorage	Sharing experiences with the community.
Network support	Communicating belonging to a group of persons with similar concerns or experiences.	Access	Offers to provide the recipient with access to new companions.
		Presence	Offers to spend time with the person, to be there.

		Companions	Reminds the person of availability of companions, of others who are similar in interests or experience.
Emotional support	Communicating love, concern, or empathy.	Relationship	Express closeness and togetherness.
		Virtual affection	Display virtual acts of affection.
		Confidentiality	Promise not to tell others. Promises to keep the recipient's problem in confidence.
		Sympathy	Expresses sorrow or regret for the recipient's situation or distress.
		Understanding and empathy	"I understand," self-disclose. Expresses understanding of the situation or discloses a personal situation that

			communicates understanding.
		Encouragement	Provides the recipient with hope and confidence.
		Prayer	Prays with the recipient.
		Expresses concern	Inquires after well-being.
		Reassurance	Nonspecific comfort.
Negative behaviors	Sarcasm, criticism, disagreement, interruption, complaint, refusals to help.	Interrupt	Changes subject or interrupt others.
		Complain	Talks about own problems.
		Criticism	Negative comments about other or blaming.
		Isolation	Will not help other, will not discuss it.
		Disagree or disapprove	Does not agree with other.
Support	Support being	N/A	N/A

requested	requested		
Support provided	Support being provided	N/A	N/A

Survivorship Care Plan Domains

Category	Definition	Example
Cancer surveillance and screening	Focuses on surveillance for recurrence of primary cancer.	Imaging (e.g., CT, MRI, mammography, colonoscopy, ultrasound) and laboratory tests (e.g., PSA, CA125).
Late effects/side effects management	Focuses on potential late or long-term side effects and their screening or management.	Bone health, cardiac dysfunction, lymphedema, sexual health/fertility, fatigue, neuropathy, cognitive dysfunction.
Risk reduction and cancer prevention	Focuses on preventive screening for	Preventive screening (gynecologic screening, colorectal screening, breast

	new cancers as well as (behavioral) risk reduction strategies.	screening, prostate screening, skin screening, lung cancer screening, and genetic screening) and risk screening (smoking cessation, weight loss, exercise/ physical activity, diet, vaccinations).
--	--	--

Psychosocial functioning	Focuses on psychosocial issues related to cancer and its treatment.	Distress, financial issues, body image concerns, social support concerns.
--------------------------	---	---

Psychosocial Need Themes

Category	Definition
Physical	Needs and issues experienced in or affecting the body, including pain, symptoms, sexual dysfunction, and care of body (such as diet, exercise, and rest).
Financial	Needs related to money, insurance, and the

	affordability of needed services and products.
Education and information	Needs related to unanswered questions and the lack of knowledge regarding what to expect as a cancer survivor, follow-up care, self-care, cancer and health research, and cancer risks, causes, and prevention.
Personal control	Needs related to an individual's ability to maintain autonomy in terms of the physical self (sexual function, evacuation, and ambulation) and the social self (disclosure about cancer and ability to make plans and socialize). Also includes wishes to return to "normal" and finding a "new normal."
System-of-care	Needs related to the health care system, including constraints and flaws that affect early detection, diagnosis, treatment, follow-up care, continuity of care, and inadequate response from health care providers.
Resources	Needs related to availability and access to supplies, equipment, therapies and medications

	(including alternative and complementary), and transportation services.
Emotional and mental health	Needs related to psychological issues, including fear (recurrence, new cancers, death, and dying), depression, anxiety, and negative feelings (mistrust toward body, anger, and guilt).
Social support	Needs related to psychosocial and interpersonal issues, including intimacy, access to support groups, opportunities to use one's own experiences to help others, and participation in social situations.
Cure	Needs related to a wish for a cure for cancer and hopes of effective treatments for self and others.
Body image	Needs related to negative perception of body, including feeling unattractive and/or ashamed and loss of trust in body.
Survivor identity	Pertains to the respondent either explicitly identifying or not identifying as a cancer

	<p>survivor because the respondent does not like the term “survivor” or feels that he or she has not reached a specific milestone to be called a survivor (e.g., not still in treatment or living a specific number of years since the diagnosis).</p>
Employment	Needs pertaining to maintaining or obtaining a source of income that is appropriate given the cancer experience.
Existential/spiritual	Needs pertaining to attaining peace in life and spirituality and making sense or meaning of the cancer experience.
Peer support*	Needs related to peer support.
School*	Needs related to returning to school.
Dating and relationships*	Needs related to dating and relationships.
Societal	Needs revealed from respondents’ commentary about conditions and issues related to society’s response to cancer, including social norms, discrimination, misinformation, policies, and resource allocation (insurance coverage).

Communication	Needs related to discourse (talking) and information exchange (explaining) about cancer and cancer experience with others (including survivor and doctor and survivor and family/friends/employers) and among medical providers.	
Provider relationship	Needs related to trust in health care providers, including decision-making, follow-through, follow-up, and support.	
Fertility/reproductive health*	Needs related to fertility and reproductive health.	
Parenthood*	Needs related to parenthood.	
Message Author		
Category	Definition	Example
Cancer survivor	A patient who has had cancer from the time of diagnosis through the	-

	<p>remainder of his</p> <p>or her life.</p>	
Family caregiver	<p>A relative who provides support on a daily or intermittent basis for an ill or disabled loved one at home.</p>	<p>Spouse, life partner, child, grandchild, grandparent, sibling, son-in-law or daughter-in-law, other relative.</p>
Informal caregiver	<p>Typically, a family friend or neighbor, provides care, typically unpaid, to someone with whom they have a personal relationship.</p> <p>This differs</p>	<p>Boyfriend or girlfriend, friend, family friend, neighbor, other non-relative.</p>

	slightly from a	
	family caregiver	
	in that an	
	informal	
	caregiver is	
	typically not	
	directly related	
	to the care	
	recipient.	
Other: CSN staff/moderator*	-	-
Other: Undiagnosed patient*	-	-
<p><i>Note.</i> Social support categories and subcategories were adapted from works by Cutrona and Suhr, and Coulson et al. (Coulson et al., 2007; Cutrona & Suhr 1992, 1994; Suhr, 1990; Suhr et al., 2004). Survivorship care plan domains were adapted from Palos and Gilmore, and Rodriguez and Lewis-Patterson (Palos & Gilmore, 2019; Rodriguez & Lewis-Patterson, 2019). Psychosocial need themes were adapted from Burg et al. (Burg et al., 2015). In addition, categories marked with an asterisk were added after the pilot test of the coding scheme. Lastly, message author categories were adapted from Griswold Home Care (Griswold Home Care, n.d.).</p>		

Appendix D

LIWC Output Variables

LIWC Categories	
Summary Language Variables	
Subcategories	Examples
Wordcount	-
Analytical thinking	-
Clout	-
Authentic	-
Emotional tone	-
Words/sentence	-
Words > 6 letters	-
Dictionary words	-
Linguistic Dimensions	
Subcategories	Examples
Total function words	it, to, no, very
Total pronouns	I, them, itself

Subcategories	Examples
Personal pronouns	I, them, her
1st pers singular	I, me, mine
1st pers plural	we, us, our
2nd person	you, your, thou
3rd pers singular	she, her, him
3rd pers plural	they, their, they'd
Impersonal pronouns	it, it's, those
Articles	a, an, the
Prepositions	to, with, above
Auxiliary verbs	am, will, have
Common Adverbs	very, really
Conjunctions	and, but, whereas
Negations	no, not, never
Other Grammar	
Subcategories	Examples
Common verbs	eat, come, carry
Common adjectives	free, happy, long
Comparisons	greater, best, after
Interrogatives	how, when, what
Numbers	second, thousand
Quantifiers	few, many, much

Psychological Processes			
Subcategories	Examples	Subcategories	Examples
Affective processes	happy, cried	Social processes	mate, talk, they
Positive emotion	love, nice, sweet	Family	daughter, dad, aunt
Negative emotion	hurt, ugly, nasty	Friends	buddy, neighbor
Anxiety	worried, fearful	Female references	girl, her, mom
Anger	hate, kill, annoyed	Male references	boy, his, dad
Sadness	crying, grief, sad		
Subcategories	Examples	Subcategories	Examples
Cognitive processes	cause, know, ought	Perceptual processes	look, heard, feeling
Insight	think, know	See	view, saw, seen
Causation	because, effect	Hear	listen, hearing
Discrepancy	should, would	Feel	feels, touch
Tentative	maybe, perhaps		
Certainty	always, never		
Differentiation	hasn't, but, else		
Subcategories	Examples	Subcategories	Examples
Biological processes	eat, blood, pain	Drives	
Body	cheek, hands, spit	Affiliation	ally, friend, social
Health Sexual	clinic, flu, pill	Achievement	win, success, better
Ingestion	horny, love, incest	Power	superior, bully
	dish, eat, pizza	Reward	take, prize, benefit

		Risk	danger, doubt
Subcategories	Examples	Subcategories	Examples
Time orientations		Relativity	area, bend, exit
Past focus	ago, did, talked	Motion	arrive, car, go
Present focus	today, is, now	Space	down, in, thin
Future focus	may, will, soon	Time	end, until, season
Subcategories	Examples	Subcategories	Examples
Personal concerns		Informal language	
Work	job, majors, xerox	Swear words	fuck, damn, shit
Leisure	cook, chat, movie	Netspeak	btw, lol, thx
Home	kitchen, landlord	Assent	agree, OK, yes
Money	audit, cash, owe	Nonfluencies	er, hm, umm
Religion	altar, church	Fillers	I mean, you know
Death	bury, coffin, kill		

Note. Adapted from: Pennebaker et al (Pennebaker et al., 2015).

Appendix E

Descriptive statistics for LIWC variables

LIWC Category	Total (SD)	Support Requested (SD)	Support Provided (SD)	<i>t</i>	<i>p</i>
Linguistic Processes					
Word count	116.85 (128.52)	147.04 (136.78)	129.65 (134.13)	2.54	0.011
Analytic	38.18 (27.11)	35.41 (23.09)	35.6 (24.83)	-0.15	0.878
Clout	52.35 (30.95)	41.02 (27.86)	53.98 (30.69)	-8.62	1.508
Authentic	63.18 (34.03)	73.49 (29.03)	64.1 (32.13)	5.97	0.000
Tone	59.42 (34.30)	48.83 (33.42)	63.8 (33.06)	-8.90	0.000
WPS	15.97 (13.69)	17.97 (16.04)	16.5 (14.35)	1.94	0.052
Words > 6 letters	13.86 (6.71)	13.82 (4.89)	13.14 (5.73)	2.46	0.013
Dictionary words	89.76 (7.51)	90.5 (5.63)	90.35 (6.77)	0.45	0.650
Function words	56.27 (8.70)	56.93 (6.84)	56.71 (7.88)	0.57	0.563
Total pronouns	17.96 (5.65)	17.97 (4.55)	18.46 (5.15)	-1.95	0.051
Personal pronouns	12.14 (5.09)	12.27 (3.88)	12.92 (4.37)	-3.04	0.002
1st person singular	7.70 (4.72)	9.08 (3.9)	7.84 (4.45)	5.74	0.000
1st person plural	0.58 (1.39)	0.48 (1.14)	0.63 (1.39)	-2.27	0.023

2nd person	2.78 (3.47)	1.51 (2.29)	3.36 (3.42)	- 11.95	0.000
3rd person singular	0.57 (1.60)	0.75 (1.89)	0.51 (1.41)	2.98	0.002
3rd person plural	0.51 (1.03)	0.45 (0.86)	0.58 (1.09)	-2.49	0.012
Impersonal pronouns	5.82 (3.85)	5.69 (3.14)	5.53 (3.52)	0.93	0.34
Articles	4.97 (3.92)	4.28 (2.45)	4.68 (2.76)	-2.93	0.003
Prepositions	12.17 (4.35)	12.77 (3.65)	12.2 (3.99)	2.90	0.003
Auxiliary verbs	11.50 (4.70)	11.52 (3.71)	11.21 (4.25)	1.52	0.126
Adverbs	6.17 (4.08)	6.39 (3.14)	6.37 (3.61)	0.06	0.951
Conjunctions	6.54 (3.42)	7.07 (2.68)	6.93 (3.19)	0.91	0.360
Negations	1.54 (1.91)	1.59 (1.57)	1.61 (1.74)	-0.30	0.757
Other Grammar					
Common verbs	19.41 (6.23)	19.51 (4.61)	19.26 (5.31)	0.99	0.317
Common adjectives	4.55 (3.55)	4.58 (2.52)	4.83 (3.15)	-1.67	0.094
Comparisons	2.09 (2.31)	2.21 (1.95)	2.26 (2.34)	-0.51	0.605
Interrogatives	1.51 (2.04)	1.57 (2.01)	1.55 (1.86)	0.24	0.809
Number	2.30 (2.85)	2.61 (2.51)	2.33 (2.81)	2.05	0.039
Quantifiers	2.23 (2.23)	2.46 (2.06)	2.26 (2.05)	1.99	0.046
Psychological Processes					
Affective processes	6.11 (5.99)	5.14 (2.95)	6.21 (4.32)	-5.47	0.000
Positive emotion	4.58 (5.84)	3.2 (2.52)	4.68 (4.09)	-8.11	0.000
Negative emotion	1.45 (1.98)	1.87 (1.82)	1.43 (1.85)	4.72	0.000
Anxiety	0.29 (0.67)	0.43 (0.78)	0.28 (0.63)	4.36	0.000
Anger	0.25 (0.78)	0.3 (0.71)	0.23 (0.67)	2.03	0.042
Sadness	0.40 (1.04)	0.49 (0.86)	0.4 (0.97)	2.10	0.035

Social processes	10.97 (6.39)	9.79 (5.61)	11.27 (6.16)	-4.89	0.000
Family	0.51 (1.04)	0.64 (1.13)	0.52 (0.99)	2.16	0.030
Friends	0.45 (1.13)	0.45 (0.89)	0.41 (0.96)	0.80	0.420
Female references	0.48 (1.40)	0.55 (1.43)	0.5 (1.42)	0.68	0.492
Male references	0.65 (1.58)	0.85 (1.89)	0.57 (1.32)	3.64	0.000
Cognitive processes	12.77 (6.27)	13.43 (5)	13.23 (5.72)	0.70	0.478
Insight	3.09 (3.06)	3.16 (2.25)	3.22 (2.84)	-0.43	0.660
Causation	1.23 (1.64)	1.33 (1.35)	1.27 (1.61)	0.75	0.447
Discrepancy	2.03 (2.18)	1.97 (1.77)	2.2 (2.22)	-2.14	0.031
Tentative	3.49 (2.96)	3.88 (2.87)	3.41 (2.58)	3.41	0.000
Certainty	1.54 (1.92)	1.5 (1.41)	1.66 (1.99)	-1.75	0.079
Differentiation	3.35 (2.74)	3.55 (2.51)	3.42 (2.44)	0.98	0.325
Perceptual processes	1.89 (2.23)	1.99 (1.73)	2.01 (2.16)	-0.21	0.825
See	0.67 (1.55)	0.65 (1.14)	0.64 (1.25)	0.15	0.874
Hear	0.46 (1.09)	0.39 (0.65)	0.56 (1.25)	-3.06	0.002
Feel	0.64 (1.10)	0.79 (1.04)	0.68 (1.14)	1.93	0.053
Biological processes	3.95 (3.47)	4.68 (3.22)	4.06 (3.38)	3.66	0.000
Body	0.78 (1.59)	0.98 (1.47)	0.81 (1.63)	2.20	0.027
Health	2.85 (2.68)	3.37 (2.44)	2.93 (2.64)	3.44	0.000
Sexuality	0.18 (0.73)	0.23 (0.77)	0.17 (0.67)	1.61	0.107
Ingestion	0.16 (0.69)	0.14 (0.52)	0.16 (0.67)	-0.44	0.654
Drives	7.69 (5.08)	7.14 (4.05)	8.08 (4.46)	-4.30	0.000
Affiliation	2.91 (3.37)	2.65 (2.93)	2.99 (3.21)	-2.15	0.030
Achievement	1.60 (1.96)	1.63 (1.74)	1.66 (1.89)	-0.29	0.765
Power	1.76 (2.22)	1.73 (1.85)	1.89 (2.3)	-1.49	0.133
Reward	1.70 (2.91)	1.28 (1.47)	1.9 (2.54)	-5.51	0.000

Risk	0.43 (1.08)	0.49 (0.9)	0.42 (1.09)	1.25	0.211
Time orientations					
Past focus	4.36 (3.48)	4.73 (3.08)	4.15 (3.27)	3.60	0.000
Present focus	13.35 (6.06)	13.29 (4.4)	13.11 (5.04)	0.76	0.446
Future focus	1.50 (2.01)	1.35 (1.55)	1.69 (1.98)	-3.76	0.000
Relativity	12.68 (5.68)	13.59 (4.99)	12.59 (5.31)	3.81	0.000
Motion	1.76 (2.29)	1.54 (1.49)	1.51 (1.69)	0.35	0.719
Space	5.16 (3.66)	5.21 (2.88)	5.34 (3.16)	-0.78	0.431
Time	5.90 (4.15)	6.98 (3.66)	5.89 (3.87)	5.66	0.000
Personal concerns					
Work	1.21 (2.03)	1.08 (1.56)	1.27 (2.03)	-2.01	0.043
Leisure	0.62 (1.39)	0.54 (1.09)	0.66 (1.36)	-1.80	0.070
Home	0.22 (0.79)	0.28 (0.91)	0.23 (0.74)	1.13	0.256
Money	0.29 (0.97)	0.23 (0.66)	0.31 (0.86)	-1.88	0.060
Religion	0.38 (3.18)	0.18 (0.55)	0.52 (3.97)	-2.09	0.036
Death	0.05 (0.28)	0.07 (0.27)	0.06 (0.29)	1.31	0.187
Informal language	1.51 (2.98)	0.9 (1.58)	1.58 (2.8)	-5.51	0.000
Swear words	0.06 (0.41)	0.06 (0.28)	0.06 (0.42)	0.25	0.798
Netspeak	0.72 (2.19)	0.38 (1.01)	0.68 (1.82)	-3.69	0.000
Assent	0.24 (0.93)	0.14 (0.79)	0.29 (0.94)	-3.23	0.001
Nonfluencies	0.26 (0.77)	0.18 (0.45)	0.31 (0.79)	-3.79	0.000
Fillers	0.05 (0.33)	0.04 (0.25)	0.06 (0.34)	-0.94	0.342
