Becoming a Dentist: Tracing Professional Identity Development through Mixed-Methods Data Mining of Student Reflections

Alyssa Wise, New York University, alyssa.wise@nyu.edu
Sameen Reza, New York University, sr4918@nyu.edu
Rujun Han, University of Southern California, rujunhan@usc.edu

Abstract: This study used mixed-methods data mining to identify concepts related to dental students' professional identity development and trace their evolution over time. Words in students' reflections were clustered based on similarity of use (represented in a 50-dimensional word-embedding model); these concepts were characterized at three points in time by constructing co-occurrence networks of constituent words and inspecting representative sentences. Analysis of 4,069 reflection statements from 378 students over four years revealed twelve relevant concepts. Over time, students' vision of *Becoming a Professional* showed increased emphasis on ethical practices, shifting focus from self-as-dentist to *Providing the Best Patient Care*. In addition to gaining *Confidence through Progress* to be a part of the *Field of Dentistry*, students also personalized their view of the profession, coming to *Know Oneself as a Dentist* and, at times, connecting this to their *Life Vision*. Implications for the theory, study and support of professional identity development are discussed.

Keywords: professional identity, word embedding, health professions education

Introduction

Development of one's professional identity is an important goal in health professions education (Cooke et al., 2010). Thus, along with "coming to know," health professions education involves an important component of "coming to be" (Wise & Cui, 2019). The importance of such disciplinary identity development, both as a form of learning and a form which shapes learning, is well-established in the learning sciences (Lee, 2017). However, while the complex processes of identifying "as-part-of" a field have been studied for young learners, little work has examined these processes for those in formal training for the professions. In health professions specifically, the work which does exist has primarily sought to measure development of professional identity, either by asking students to rate themselves (Buck et al., 2019) or by grading students' responses to specific scenarios (Kalet et al., 2018). Such quantitative approaches have not resulted in consistent findings about if, or when, professional identity development occurs, nor help unpack why differing results were found - an important task for research given evidence that students develop professional identity differently (Vivekananda-Schmidt et al., 2015).

In contrast, texts students write about authentic professional situations they face offer a rich source of data for providing insight into identity development from the bottom up. Reflection is an important activity in health professions education, emphasized as a way to help students articulate and learn from episodes in their academic journey, recognize significant moments, and develop strategies towards future goals (Chirema, 2007). Written on an ongoing basis throughout students' program, these texts provide a unique opportunity to trace the development of students' ideas and concepts about their learning and themselves over time. While some initial efforts to examine reflections for insight into professional identity development have been made, these studies were limited either in scope (by conducting manual analysis on small samples of data, Jonas-Dwyer et al., 2013) or depth (by applying computational topic models to identify general subjects being discussed, Chen et al., 2016).

The current study addresses these limitations by applying a mixed-methods data-mining approach to identify key concepts for students' professional identities and how they change over time. The work is conducted in the case of dentistry education. Results contribute to an understanding of the critical concepts related to dental students' professional identity development and efforts to support it. In addition, the methodological approach can be applied to reflections in other professional contexts to study similarities and differences across fields.

Conceptualizing and studying professional identity development

At a basic level, professional identity is defined as feeling part of an identified group (Vivekananda-Schmidt et al., 2015). It includes possessing "attitudes, values, knowledge, beliefs and skills" (Adams et al., 2006, p. 56) shared within this group. Inherent in this definition is both the internal state of an individual and the ways this shapes their relationship to a larger community (Carlone, 2017). Importantly, the internal state is not simply a set

of beliefs, values, norms and aspirations but the integration of these qualities into one's sense of self. In this way, professional identity differs from simple "outward professionalism" (how one *acts* in the world) to the internalization of these attitudes to form part of one's self-concept (how one *is* in the world, Bebeau & Monson, 2012). Put together, these elements are captured in Cruess et al.'s (2014) definition of professional identity as "a representation of self, achieved…over time, during which the characteristics, values and norms of the […] profession are internalized resulting in an individual thinking, acting and feeling like a […]" (p. 1447).

There are two broad ways in which the development of professional identity in the health professions has been conceptualized. In the first, professional identity development is defined as a closer integration between one's perception of self and one's "perception of future professional role" (Vivekananda-Schmidt et al., 2015, p. 15). Here the emphasis is on closing the gap between how one views oneself (one's "moral identity", Bebeau & Monson, 2012, pg. 136) and how one sees members of the group to which they aspire to belong (the core values of a profession, Buck et al., 2019). Different professional experiences can reduce (or increase) this gap, with development defined as overall progress towards becoming one with the expected role. A mature professional identity is thus well-aligned with respect to the questions: "Who am I as a professional?" and "Who do I aspire to become?". The second conceptualization of professional identity development takes a more structured approach, requiring a person to move through a series of identity stages. One model within this conceptualization (Cruess et al., 2016) defines five stages starting from an initial state that is based on individual characteristics and prior work experience (Adams et al., 2006). As students learns about their future role, they develop both knowledge of the behavioral norms expected of them and an understanding of when particular behaviors are appropriate; this is followed by the ability to demonstrate these behaviors under supervision and then consciously perform them on one's own. The progression from "knows" to "knows how" and then from "shows how" to "does" culminates in an identity of "is" when one consistently demonstrates not only the behaviors, but also the values and attitudes expected of them in their professional role (Cruess et al., 2016). While this model focuses primarily on the experience of an individual, other stage-based conceptualization models include greater attention to identity development through collaborative work in professional teams. For example, Bebeau & Lewis (2003) presents a four-stage model of professional identity formation for healthcare professions that proceeds from Initiation to Independent Operator to Team Oriented Idealist to Self-Defining Professional.

While the two conceptualizations differ in some ways, they are consistent in seeing development as episodic (as opposed to continual), triggered by professional experiences, and facilitated by reflection on those experiences (Vivekananda-Schmidt et al., 2015). While there is often a presumption that as students progress to more advanced stages of their educational journey their professional identity will naturally develop, empirical work to date has failed to consistently document such patterns (Crossley and Vivekananda-Schmidt, 2009; Buck et al., 2019; Kalet et al., 2018). For example, using a five-point scale Buck et al., (2019) found only a slight shift in the distribution of students across eight medicals schools showing higher levels of self-reported Physician Professional Identity in clinical years 3 and 4 than in pre-clinical years 1 and 2. Levels of Identity Integration scores were stable across the three initial years but declined in year 4. This unexpected decline may be due to students facing "identity dissonance" (Cruess et al., 2014, p. 1448) for the first time as they are given more opportunities to participate in real medical situations yet may not be able to confidently handle unpredictable problems that arise. In contrast, comparing responses from 1,334 student doctors and dentists on their feelings of competence in nine areas (e.g. emergencies, teamwork) Crossley and Schmidt (2009) found that students in later years showed higher mean scores than those in earlier years. Moving from cross-sectional to longitudinal designs, Kalet (2018), assessed 130 students twice over fifteen months using Bebeau & Lewis's (2003) four-stage model applied to Professional Identity Essays (Bebeau & Monson, 2012). Results showed half of the students remained in the same stage while the other half progressed, but only by half a stage.

The above review highlights limitations of past work to study professional identity. First, measurement schemes that offer limited possible levels may be too coarse to trace the slow process of identity development which can continue throughout one's career. Second, even with an appropriate scale, numerical assessment places students along a development continuum but provides no information about how the process occurs, what are different aspects of identity, or if these develop in the same ways. Trede et al., (2012) have pointed to the lack of explicit conceptualization of facets of professional identity as an impediment to understanding its development. Finally, student responses to designed situations focus on aspects of identity development important to researchers, which may differ from those valued by students themselves, limiting their ecological validity.

The current study
The current study addresses the above limitations by taking advantage of an existing, authentic data source of student reflections, applying mixed-methods data mining to identify key concepts in how students think about becoming a healthcare professional and how they change over time. Analytics (specifically word embedding,

clustering and network analysis) are used to extract key concepts from the reflection texts, operationalized as sets of words that are used in similar ways and their co-occurrence. Concepts (networks of words) created at multiple points in time are analyzed both quantitatively and used as a "way in" to qualitatively examine specific text from the larger corpus. This approach allows for rich examination of a large data corpus to generate understanding of the process of identity development. We ask two questions. RQ1: What are key concepts related to dental students' professional identity displayed in their reflection statements? RQ2: How do these concepts change over time?

Methods

Learning context and data

Dental students enrolled in a large US university are required to reflect on their courses, competencies, knowledge acquisition and professional development on an ongoing basis throughout their four-year program. Reflections are written and stored in a purposefully designed online system. All data from a single graduating class over their four years of study (392 students; 284,320 reflections) were extracted from the system. The main target of analysis was a sample taken from a subset of the 12,564 reflections focused on professional development (Figure 1).
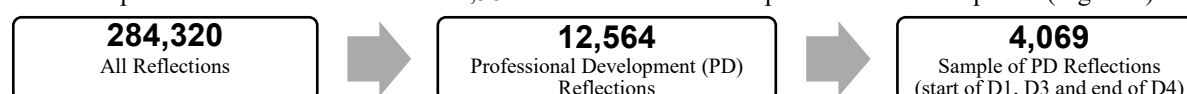

| **284,320** All Reflections | → | **12,564** Professional Development (PD) Reflections | → | **4,069** Sample of PD Reflections (start of D1, D3 and end of D4) |

Figure 1. Sampling Process for Analysis.

Key concept identification: Word embedding model construction and clustering

Word embedding models represent the vocabulary of a corpus as vectors in a multi-dimensional space based on how they are used in context. Such models require extremely large corpuses to train, thus the complete set of 284,320 reflections (6,752,282 instances of 11,604 unique words) was used for this initial phase of analysis. FastText (Mikolov et al., 2013), an efficient library for performing learning of word representation tasks, was used to build a 50-dimensional vector representation of the 11,604 unique words. Distance between words in the vector space is a proxy for similarity of use; clustering produces sets of words that can be considered together as a concept. As students reflected more in their final years, the corpus and model were weighted towards similarities in this period; this allowed us to trace concepts that maintained or developed some importance over time but may have obscured concepts important only at the start. A histogram was used to select a subset of frequently used words for clustering as inputting many thousands of words produces complex and uninterpretable output. Large declines in word usage were observed after the 29th, 39th, 51st, 60th, 77th and 127th words. 127 words was selected as a set concise enough to allow clustering while also large enough to represent the breadth of key concepts commonly discussed. Agglomerative clustering (Ward's method, Euclidean distance) was applied and a scree plot indicated 13, 17 and 18 as potential solutions. While additional clusters can sometimes unnecessarily divide coherent groupings, in this case the larger N solutions provided finer distinctions that supported interpretability, thus the 18-cluster solution was selected. Cluster (concept) size ranged from 2 to 18 words.

Characterizing concepts based on network structure

The remaining analysis focused on students' professional development reflections (six prompts at start and end of each year). Reflections were sampled at the beginning (start of year one "D1" 1,020 reflections) middle (start of year 3 "D3" 940 reflections) and conclusion (end of year 4 "D4" 2,109 reflections) of the program, yielding 4,069 reflections from 378 students. Networks was created for each concept and point in time following Wise and Cui (2019). Constituent words in a concept were represented as nodes (size indicated sentence-based frequency); edges between nodes were based on the number of sentences in which the two words co-occurred. Networks were examined to characterize each concept in terms of most prevalent words and the strength of their connections. Changes in concept prevalence over time were tracked based on the percent of sentences in which any of the constituent words appeared; relative comparisons across concepts were made based on the percent of sentences in which either of the top two constituent words appeared to account for varying numbers of words in concepts.

Characterizing concepts based on representative sentences

One hundred and fifty sentences were sampled for each of the 18 concepts at each point in time. Network structures were used to frame sampling by restricting the extractable corpus to sentences which contained multiple constituent words of a concept. Sampled sentences were ordered based on prevalence of word pair co-occurrence to prioritize understanding the nature of the strongest connections in the network. The constant-comparative method (Gibson & Brown, 2009) was used to identify emerging themes, initially within each concept at each time

point, subsequently over time, and finally across concepts. The goal was to describe each concept, identifying if, and how, it changed over time; concepts labels were developed to reflect key aspects of each characterization.

Findings

A total of eighteen concepts were identified in the student reflections examined. Six were concepts about the process of dental education that did not relate to identity development and were excluded from further analysis. The remaining twelve concepts related to four different areas (Table 1). *(a) Becoming a Professional* and *(f) Providing Best Patient Care* were the most frequent concepts across all time points, with the latter increasing in D3 and D4. Several concepts also showed changes in the relative prominence of particular words, strength of connections between them, or the specific ways the words were used in context

Table 1: Sentence-Based Frequency of Concepts over Time + Relative Prevalence in D4 (based on top 2 words)

	D1 (start)	D3 (start)	D4 (end)	Rel. Prev.
Concepts Related to Competence and Professionalism				
(a) **Becoming a Professional:** professional, ethical, believe	1353 (36%)	1438(35%)	4492(36%)	High
(b) **Life Vision:** life, long, personal	251(7%)	185(4%)	683(5%)	Low
(c) **Field of Dentistry:** field, career, profession	609(16%)	476(11%)	1706(14%)	Med
(d) **Become a Competent Dentist:** become, competent, dentist	984(26%)	713(17%)	2026(16%)	Med
(e) **Gain Knowledge and Skills:** gain, skills, knowledge, clinical	736(20%)	627(15%)	1867(15%)	Med
Concepts Related to What it Means to be a Dentist				
(f) **Providing Best Patient Care:** patients, care, possible, needs	1213(32%)	1576(38%)	4676(37%)	High
(g) **Responsibility as Professionals:** trust, must	512(14%)	483(12%)	1415(11%)	Low
(h) **Promoting Oral Health:** oral, health	281(7%)	261(6%)	814(6%)	Low
Concepts Related to Developing Self Awareness				
(i) **Confidence Through Progress:** time, feel, learned, progress	623(17%)	1063(26%)	3552(28%)	Med
(j) **Knowing Oneself as a Dentist:** would, know, really, like	1185(31%)	1439(35%)	4125(33%)	Med
Concepts Related to Community and Culture				
(k) **Becoming Part of Community:** people, community, others	559 (15%)	379(9%)	1044(8%)	Low
(l) **Collaborative Environment:** faculty, peers, colleagues	284(8%)	249(6%)	679(5%)	Low

Note: % indicates proportion of total sentences submitted in that time period (D1=3772, D3=4160, D4=12,531)

Concepts related to competence and professionalism

The most prevalent concept related to competence and professionalism offered a broad view on what it meant to become a member of the dental profession. **Becoming a Professional** (Figure 2a) characterized students' vision of coming to be a dental professional, including academic, ethical and moral dimensions. As students progressed from D1 to D4 there was a shift from the construction of becoming a professional as upholding a set of general standards *"**Becoming a professional** requires excellent practical skill as well as knowledge, societal responsibility, sincere concern, compassion for others and willingness to accept new ideas for improvement"* (D1) to an increasingly concrete and personal view on what students thought becoming a dental professional meant to them *"However, I **believe** the most **important** component of being a **professional** is taking responsibility for our patients and try to respect and understand patients' point of view without compromising to provide standard treatment care"* (D3) as well as an increase in connection to the ethical dimensions of practice *"During my residency next year I plan on further establishing my competence in being an **ethical professional**"* (D4).

An additional concept in this group, **Life Vision** (Figure 2b), also related to the notion of a personal view on professional life expressed students' long-term life goals to be able to successfully incorporate multiple facets of their personal and professional lives together. This concept was used relatively infrequently across all years, but appeared to represent an important part of identity formation for some students, by expressing values they wished to uphold throughout their lives: *"I have a strong passion for social service, and I strive to incorporate it not only in my **personal life** but work **life** as well."*
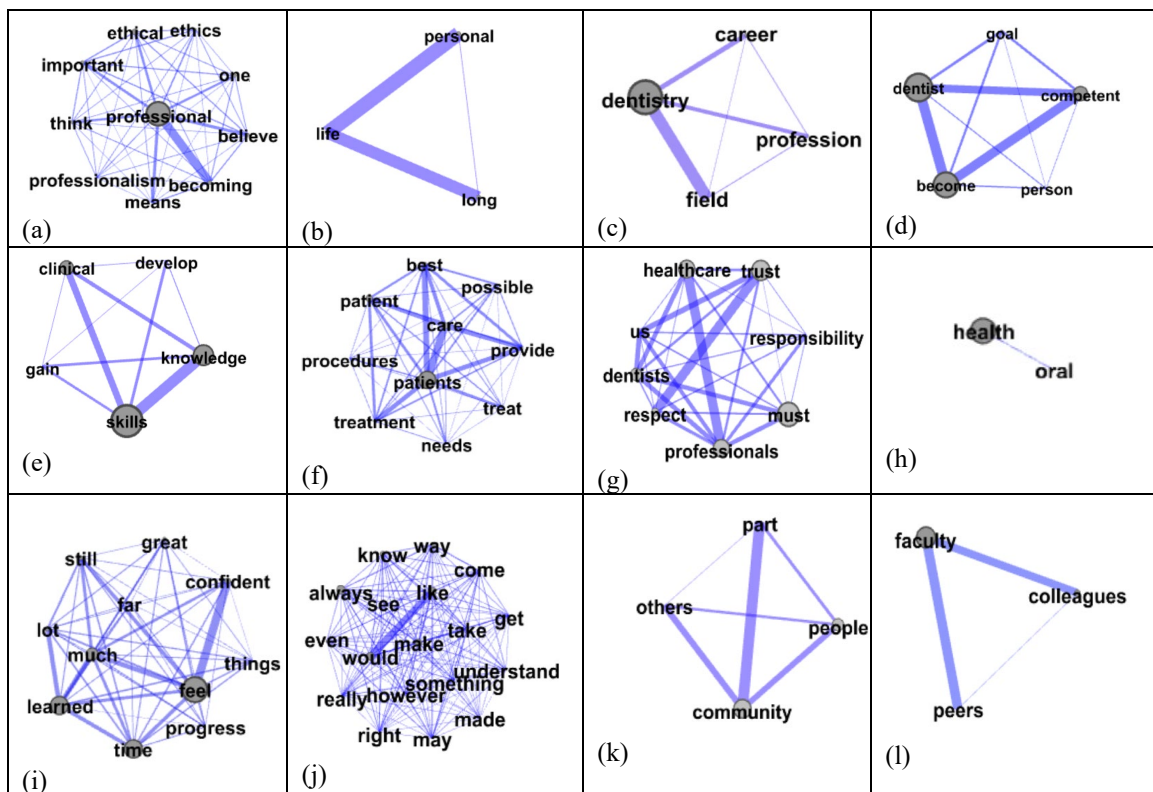
Figure 2. Networks for concepts (a - l) in D4 following key in Table 1.

The next several concepts existed as a complementary trio. **Field of Dentistry** (Figure 2c) represented students' view of dentistry as a challenging field that demands continuous learning to keep abreast of the latest treatments. This field was presented by students as apart from themselves, constructed as a set of normative behaviors that "one" performs. For example, *"[working]…in the **field** of **dentistry** entails showing up to appointments on time, treating patients with respect, treating peers with respect, and maintaining the ability to provide competent care"* (D4). **Become a Competent Dentist** (Figure 2d) then articulated students' personal objectives to meet this external standard by becoming proficient in the required activities. For example, *"I would…like to **become** a **competent** and skillful **dentist**, I want to be proficient, both in procedure and patient interaction so I can develop meaningful patient relationships"* (D4). Statements of **Become a Competent Dentist** declined in prevalence from D1 to D3 and D4 (Table 1) and later mention of this concept acknowledged progress made. Finally, **Gain Skills and Knowledge** (Figure 2e) offered specification of how students hoped to achieve competency by focusing on the specific knowledge and skills they felt necessary to acquire. For example, *"I wish to **gain** the practical **skills** in performing procedures such as implants, root canals, and complicated tooth extractions"* (D4). Skills and knowledge included technical elements as well as those related to patient communication, time management and organization, offering a view into the full complement of aptitudes students saw as required for competent practice. Later reflections emphasized the use of knowledge and skills in clinical situations.

Concepts related to what it means to be a dentist

Providing Best Patient Care (Figure 2f) was a prevalent concept at all points in time, but also showed a sharp rise in D3 that was maintained in D4 (Table 1). It described students' intention to serve patients' needs with the best procedures and care possible. While earlier reflections (D1, D3) focused more strongly on what the students could do to *provide* such *care* : *"[my goal is to] acquire the skills to provide excellent quality **care** to **patients** as a well-rounded confident dentist"* (D1), over time the concept becomes increasingly centered around *patient(s)*, their *needs* and *possible* ways to meet them *"It is very important to understand **patients'** concerns, **needs**, and expectations in order to **provide** a satisfactory result without compromising much of the ideal **treatment** objectives"* (D4).

Responsibility as Professionals (Figure 2g) described the incumbent responsibility of health professionals to establish *trust*, and mutual *respect* with patients and among themselves. While the concept was

mentioned only infrequently, it showed an important shift over time from being discussed in a detached way of something that "they" as professionals should do, to being owned as something that "we" as professionals will do with time. The reflection statements show this developing sense of collective identity as use of pronouns shifted from the third person *"dentists should **be seen** as part of an ecosystem of healthcare professionals which include dentists doctors nurses and nutritionists who all work together for the betterment of **their** patients"* (D1) to the first person *"**We** won't just be titled as dentists in dental office, but society will see **us** as healthcare professionals so **we** must act the part"* (D4). The final concept related to what is means to be a dentist **Promote Oral Health** (Figure 2h) was a two-word concept (oral, health) focused on the responsibility of dentists to not only treat problems but also educate patients about taking preventative care of their mouth and teeth. *"Through my dental experience in "college" I have learned how to emphasize on educating and promoting **oral health** to every patient that I treat, since I feel that this is the duty of a dental professional"* (D4). While not very prevalent, this concept offers an additional element of dental professional identity showing the position of a dentist not just in treating problems but also in encouraging larger oral health habits.

Concepts related to developing self-awareness
Confidence through Progress (Figure 2i) described students' progress to start to *feel confident* in their growing abilities based on how much they had *learned*. This concept shows a sharp rise in D3 that was maintained in D4 (Table 1) as students describe how far they had come since they entered dental school, *"Professional **progress** is becoming more exciting and I **feel** more **confident** about certain procedures, for example, I was terrified about endodontic treatments because I don't think we got enough pre clinical experience, however with endo faculty help I **feel** more **confident** during these procedures"* (D3). **Knowing Oneself as a Dentist** (Figure 2j) described students' developing self-perception of themselves as dental practitioners, a process in which they recognize what they have *come* to *know* and *understand* through their studies. Often these were qualities of themselves, their patient handling, or choices to be *made* rather than specific dentistry skills - *"I have been able to think through these situations and **come** to **understand** that when it comes to health care there should not be any gray zone in ethics."* Often, this concept also includes aspirational statements about the kind of dentist they *would like* to be.

Concepts related to community and culture
Part of a Community (Figure 2k) described the ways that students saw themselves as playing a role in their student, professional or local community. Though the concept was low in prevalence and declined over time (Table 1), there was also an expansion in focus: in D1 students primarily mentioned serving their local community whereas in D3 and D4 they also refer to the larger dental community whose respect they strive to gain, *"Being **part** of the dental profession gives me a strong sense of pride and professional **community**, therefore I will strive to participate in as many professional meetings and conferences as possible to give back to this **community** as well as to keep up with the latest advancements."* In contrast with the broad notion of community, **Collaborative Environment** (Figure 2l) described the specific ways students worked as part of teams with peers, colleagues and faculty to serve patients *"I have learned how to brainstorm ideas with my **colleagues** and my **faculty** over complicated treatment plans, and have learned to work with dental hygienists and dental assistants so that we can more efficiently provide the best care for our patients"* (D4). This stands in contrast to the idea of dentists working as isolated professionals and more as a team of healthcare providers.

Discussion and conclusion

Contributions to understanding dentists' professional identity and its development
Put together, the evolution of the different concepts described above show a shift in dental students' identity from the start to the end of their four-year program in which they start to see themselves as becoming part of their chosen field of dentistry. At a global level, their view of this process aligns with the "closing the gap" conceptualization (Buck et al., 2019) in which an individual seeks to align their image of self with that of their future professional role (Vivekananda-Schmidt et al., 2015). This is seen in the way that student conceptualize their profession apart from themselves as the *'field of dentistry'*, then setting goals and recognizing progress to *gain the knowledge and skills* needed to *become a competent dentist*. This progress also brings about increased confidence in their own abilities to conduct certain procedures and face complex situations. In addition, over time students showed an increased emphasis on ethical practices as a component of *becoming a professional* and a shift from focus on themselves as healthcare providers to the people which they will serve; this evidence points towards adoption of a core value in dentistry i.e., patient-centered care. Through the above processes, students start to see themselves as becoming part of their desired professional group (dentists) shifting pronoun references from "they" to "we," as also seen previously in Wise and Cui (2019). This starting to see themselves "as-part-of" refers not

only to the global profession but also to local collaborative teams, suggesting alignment with the advanced "team-oriented idealist" stage of professional identity development in Bebeau and Lewis's (2003) model. Finally, along with the process of aligning themselves with what it means to become a dental professional, students also personalized their vision of the profession, expressing a vision of what kind of dentist they wanted to be and, in some cases, connected their professional identity with their personal one. This is a novel finding suggesting the need for a reconceptualization of professional identity development from simply aligning oneself with an existing profession to finding (and defining) one's place within it. From a consensus view, this might be seen as individuals sharing a core nominal identity but practicing it in individualized manner in their lives (Jenkins, 2014); or from a critical perspective, individuals may form identities "that conform to, resist, and sometimes transform meanings implied by prevailing structures" (Carlone, 2017, p. 527).

Implications for studying and supporting the development of professional identity

While the intent of this study was not to quantify growth in students' professional identity, the rise in prominence of a core value of the dental profession (a patient-centered orientation), coupled with increasing confidence, and coming to see oneself as part of the field align with theoretical conceptions of professional identity development. They also suggest that if professional identity formation needs to be assessed in a quantitative manner then the multi-dimensional markers identified in this study could provide a richer picture than the single-dimensional scales used previously. Importantly, this would not only give educators a better way to determine if development is taking place but also potentially offer insight into patterns by which different aspects of identity develop. In terms of support for identity formation, knowing that professional identity contains multiple aspects suggests different ways that support can be leveraged. For example, the reflective activities students are asked to engage in could be modified to specifically support students' engagement with issues such as "finding one's place in the field." In addition, while collaboration and community did emerge as concepts in this study, they were mentioned less frequently than many other concepts. Given the theoretical role that professional socialization is thought to play in supporting identity formation (Trede et al., 2012; Cruess et al., 2014), additional attention might be paid to incorporating social interaction into learning activities and fostering reflection on those experiences. Finally, there is the potential to build on the current work to eventually offer personalized feedback to students that can further aid in their professional identity development.

Methodological advances, limitations and future work

This study used a mixed-methods data mining approach to identify concepts related to development of dental students' professional identity by clustering words that were used similarly in reflections (indexed by their position in a 50-dimensional word embedding space). Construction of networks exposed the structure of each concept and were used as a "way in" to qualitative examination of representative sentences. This approach addressed limitations of prior work that *only* looked at abstracted representations of reflections (disregarding contextual-use of language) or conducted in-depth analysis (but only on a small data corpus). The resulting insight into the development of professional identity further establishes the value for the learning sciences of mixing data mining and qualitative methods to achieve rich analysis of big data (Wise & Cui, 2018). We also highlight the particular value of word embedding models to identify similarity and difference in words based on their use-in-context rather than morphological structure. In the current study this led to differentiation between single and plural forms of the word dentist—one part of the individually-oriented concept "Become a Competent Dentist," the other part of the community-oriented concept "Responsibility as Professionals"—a distinction that would have been lost via stemming or lemmatizing. There are, however, limitations to the current work. First, word embedding models require very large corpuses of data; thus words from all four years of reflection text data (in which later years were more heavily represented) were combined to train a single model. While the evolution of concepts was examined, changes to the constituent words of each concept were not possible and thus there are variations in concepts that may not have been detected. In addition, concepts important only at the start are unlikely to be detected. Second, this approach aggregated reflections for each concept at each point in time into a single network, precluding the detection of differences using traditional statistical methods which compare distributions of cases. Third, this approach also aggregated across students which is useful for understanding the process of professional development overall, however, does not necessarily represent the specific experiences of any particular student. Finally, while student reflections are an authentic data source, they are a required task and one which students know can be read by instructors. Thus, the comments can be perfunctory or reflect a social desirability bias.

Future research can explore the application of this mixed-methods data-mining approach to other texts that may offer indications of professional identity development for dentists or other professions and expand the current approach to allow for examination of trajectories of identity development by individual students. In addition, work is needed to explore ways in which such analyses can be used not only understand the process of

identity development globally, but also support it in specific instances. For example, what kinds of reports could be provided to faculty members, academic advisors, or students themselves to help inform the development of future dentists' professional identity?

Conclusion

This study used a mixed-methods data-mining approach to generate insight into key concepts of dental students' professional identity development. The findings help address Trede et al.,'s (2012) call for greater explicit conceptualization of the facets of professional identity as a precondition to effective work to foster its development. Beyond the specific implications of the work for the future study and support of identity development for dental (and potentially other healthcare) professionals, it also serves as an example of the value of integrating computational and qualitative analyses within the learning sciences.

References

Adams, K., Hean, S., Sturgis, P., & Clark, J. M. (2006). Investigating the factors influencing professional identity of first-year health and social care students. *Learning in Health and Social Care, 5*(2), 55-68.

Bebeau, M. J., & Lewis, P. (2003). Manual for assessing and promoting identity formation. *Minneapolis: Center for the Study of Ethical Development, University of Minnesota.*

Bebeau, M. J., & Monson, V. E. (2012). Professional identity formation and transformation across the life span. In *Learning trajectories, innovation and identity for professional development* (pp. 135-162). Springer.

Buck, E., West, C., Graham, L., Frye, A. W., & Teal, C. R. (2019). Challenges to assessing professional identity in medical students: a tale of two measures. *Medical Education Online*, *24*(1), 1649571.

Carlone, H. B. (2017). Disciplinary identity as analytic construct and design goal: Making learning sciences matter. *Journal of the Learning Sciences, 26*(3), 525-531.

Chen, Y., Yu, B., Zhang, X., & Yu, Y. (2016). Topic modeling for evaluating students' reflective writing: A case study of pre-service teachers' journals. *Proceedings of LAK'16* (pp. 1-5). Edinburgh, UK: ACM.

Chirema, K. D. (2007). The use of reflective journals in the promotion of reflection and learning in post-registration nursing students. *Nurse Education Today, 27*(3), 192-202.

Cooke, M., Irby, D. M., & O'Brien, B. C. (2010). *Educating physicians: A call for reform of medical school and residency.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Crossley, J., & Vivekananda-Schmidt, P. (2009). The development and evaluation of a Professional Self Identity Questionnaire to measure evolving professional self-identity. *Medical Teacher*, *31*(12), e603-607.

Cruess, R. L., Cruess, S. R., Boudreau, J. D., Snell, L., & Steinert, Y. (2014). Reframing medical education to support professional identity formation. *Academic Medicine, 89*(11), 1446-1451.

Cruess, R. L., Cruess, S. R., & Steinert, Y. (Eds.). (2016). *Teaching medical professionalism: supporting the development of a professional identity*. Cambridge, UK: Cambridge University Press.

Cui, Y., Wise, A. F., & Allen, K. L. (2019). Developing reflection analytics for health professions education. *Computers in Human Behavior, 100*, 305-324.

Gibson, W., & Brown, A. (2009). Working with qualitative data. Los Angeles, CA: Sage.

Jenkins, R. (2014). *Social identity*. New York: Routledge.

Jonas-Dwyer, D. R. D., Abbott, P. V., & Boyd, N. (2013). First reflections: third-year dentistry students' introduction to reflective practice. *European Journal of Dental Education*, *17*(1), e64-69.

Kalet, A., et al. (2018). Professional identity formation in medical school: One measure reflects changes during pre-clerkship training. *MedEdPublish*, *7*, 1-16.

Lee, C. D. (2017). Expanding visions of how people learn: The centrality of identity repertoires. *Journal of the Learning Sciences, 26*(3), 517-524.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781.*

Trede, F., Macklin, R., & Bridges, D. (2012). Professional identity development: A review of the higher education literature. *Studies in Higher Education*, *37*(3), 365-384.

Vivekananda-Schmidt, P., Crossley, J., & Murdoch-Eaton, D. (2015). A model of professional self-identity formation in student doctors and dentists: a mixed method study. *BMC Medical Education*, *15*(1), 83.

Wise, A. F., & Cui, Y. (2019). Top concept networks of professional education reflections. In *Proceedings of LAK'19* (pp. 260-264). Tempe, AZ: ACM.

Wise, A., & Cui, Y. (2018). Envisioning a learning analytics for the learning sciences. *Proceedings of ICLS 2018* (pp. 1799-1806). London, UK: ISLS.