Understanding the enigma that is résumé layouts by analyzing saccade-derived and scanpath metrics

Shannon Devlin
Clemson University
Sdevlin@clemson.edu

Madison Sauls
Clemson University
msauls@clemson.edu

Lisa Hieftje-Conley
Clemson University
lhieftj@clemson.edu

Alex Ward
Clemson University
atward@clemson.edu

ABSTRACT

Constructing the best résumé is a topic across all industries and business professionals. Initial investigation into résumé review highlights how brief the process is and how influential it is on a candidate's chance of selection. Previous work has explored how varying résumé characteristics influence résumé review and candidate selection with tools such as surveys and interviews, but explicitly studying the reviewer's scanning process has been limited. Previous work has shown how visual hierarchies influence information retention, usability, and a person's holistic understanding of a system. Specifically eye tracking studies reveal how saccade-derived metrics have been used to quantify how scan strategies differ on interfaces and websites, but this specific investigation with résumé has been limited. Therefore, this work investigates how visual hierarchies affect number of saccades, average saccade amplitude, total scan path length, and transition entropy as a person reviews a résumé and makes the infamous fit/no fit decision and recalls the candidate's résumé review criteria. Additionally, reviewer specific analysis is conducted to further investigate individual differences. Total scanpath length and average saccade amplitude were significantly different across visual hierarchies, transition entropy indicated transitioning between résumé review criteria was not particularly random, and scan strategies varied across reviewers. This work contributes to the résumé construction process as it highlights the role a visual hierarchy has on the résumé review process. It also adds to the understanding and application of saccade-derived eye tracking metrics.

KEYWORDS

Eye-tracking, résumé construction, saccade-derived metrics, visual hierarchy, entropy

INTRODUCTION

Current eye-tracking data shows that job recruiters may take as little as six seconds to review a résumé before deciding if the candidate will move on to the next stage of the hiring process [3,13]. This decision is the first and sometimes ultimate judgment of a job candidate. A previous study [3] found résumé reviewers spent 80% of their time reviewing a very specific list of items of information, such as previous position and education. Similar studies have identified a list of criteria that closely matches those information items [13]. These items are referred to as the résumé review criteria, and it is clear that among job recruiters and hiring managers, the search for information is not random, but targeted towards these certain résumé criteria. Although the quality of the job candidate is the main factor in the hiring decision, the physical properties of the résumé, like the location of the résumé review criteria, may also play a role, especially when reviewers are reviewing countless of résumés that are similar in nature. Any simple search engine query will produce endless amounts of recommendations on how to construct a résumé, however there is a lack of objective data to support the benefits of each format and a lot of contradictory advice [13]. Therefore, it is important to understand how the résumés' organization influences scanning patterns and thus the fit/no fit decision making process made by job recruiters.

BACKGROUND

Working memory is a major supporting factor for how individuals are able to process the constant flow of new information on a daily basis. The main components of working memory allow people to take in new information, manipulate it by either comparing it to previous knowledge or the status of the current environment, and make a decision or take an action [1]. These processes map directly onto the act of quickly reviewing a résumé, where a recruiter acquires the information visually in the form of text, compares the information to previous résumés, matches the information to predetermined qualifications, and makes the fit/no-fit decision. Therefore, working memory is likely an important cognitive process that is active during a typical résumé review.

Previous work has shown cognitive processes can be linked to specific eye tracking metrics [14]. Two common metrics are fixations and saccades where fixations are time when the eye is relatively still, and saccades are the eye transitions between those fixations [12]. Specifically, eye tracking has been used to understand the effects of information's visual configuration to attention and retention of information [4]. One study found that the efficient functioning of visual short-term memory, a subsidiary system of working memory, relies on the spatial configuration of information [10]. Specifically, information related to other pieces of information were more likely to be encoded into visual short term memory when grouped closely together in the visual field. This is important because stimuli cannot be further processed if they do not make it from initial sensory systems (i.e. information simply presented in the visual field) to working memory (i.e. actually being processed). Therefore, information presented on a résumé should be presented

strategically, so it easily reaches recruiter's working memory. This is especially true since the typical amount of time spent viewing the résumé is very limited, which subsequently means the recruiter's cognitive resources are just as limited.

Additionally, the visual layout of an item can influence eye tracking metrics related to scan patterns [2]. Previous work found a complex visual hierarchy, which is defined as content employing systematic formatting and layout organization of its headers, text, and pictures, can lead users to scan more of the webpage rather than just employing the traditional F-shaped scan pattern traditionally found in website viewing [9]. Since it is has been previously determined a résumé should follow some kind of visual hierarchy [3], studying the scan patterns across different visual hierarchies of résumés is of interest.

While research specifically tailored to résumés is limited, studies investigating saccade-derived metrics on various website types may be insightful in understanding how people review visual information in varying structures. For example, an eye tracking study had people scan various website layouts (e.g. homepages versus information pages, search engine sites versus retail sites, etc.) in order to understand differences in scanning strategies when scanning for recall versus free will [11]. In addition to investigating how fixation-derived metrics such as fixation duration and rate of fixations, saccade-derived metrics like saccade rate (number of saccades across the web page), and scanpaths using a string-editing method (i.e. the sequence of fixations on predetermined areas of interest (AOI) on a given webpage for each participant). In order to understand the variability of scanpaths across individuals, comparisons on individual scanpaths based on website type and page were completed. Saccade rate results for search engine websites indicated the page with the output list of results from a search query required more cognitive effort than its first page (where the query was entered) even though previous eye tracking metrics (e.g. number of fixation) did not indicate this. This same contradiction in fixation rate and saccade rate results was also apparent in business websites. The authors explained this finding relates to mental workload being multidimensional, and that constructs like visual complexity can contribute to website scanning pattern [11]. This work highlighted the importance in investigating saccade-derived metrics when trying to understanding the cognitive efforts associated with visual information acquisition and implementation.

Saccade-derived metrics analyze the scanning methods employed by individuals [14]. Two of these specific metrics include number of saccades and average saccade amplitude. Number of saccade have been found to indicate the amount of searching a person is doing. Previous work found poorly designed interfaces led to greater increases in total number of saccades [4]. Average saccade amplitude has been found to quantify how meaningful or salient an item is on an interface, with a larger value indicating more meaningfulness [5, 12]. They also have been found to measure how pre-planned certain eye movements are, with larger values indicating more pre-planning [5]. Knowing the amount of pre-planning a person does relates to how well the design and layout matches the person's internal representation of the system [5].

Scanpath is defined as the complete sequential saccade-fixation-saccade process a person completes when viewing a stimulus [12]. Previous evidence indicates scanpath length can measure the efficiency of someone's search, with a larger value indicating less efficient search [12]. For example, scanpath length decreased as organization to the visual information increased [4, 9]. Other metrics related to scanpath include transitions matrices, which find the probability of transitioning to and from AOIs [12, 14]. If the transition probabilities (i.e. the individual probabilities of transitioning to and from AOIs) are framed in a discrete-time Markovian chain (i.e., the probability transitioning to state *j* from state *i* only depends on currently being in state *i*) then a metric called entropy can be calculated [8]. Specifically, stationary entropy (H_s) quantifies how distributed a person's visual attention is during the task [8], and transition entropy (H_t), which measures the randomness of a person's eye movements during the task can be computed and compared [7, 8]. Previous studies found entropy values can quantify how varying ages of children with and without autism spectrum disorder view facial images [15] and has also been used when comparing individual differences in curiosity when it came to viewing different styles of art, with more curious individuals having significantly higher transition entropy values [8]. Increases in entropy were also found with increases in sentence length, location position of items, and comprehension difficulty, which may have direct effects of entropy values as a reviewer scans a résumé [6].

From on the previous findings, it is likely that résumé reviewing behavior may vary depending on the visual hierarchy. Furthermore, due to the time limitations this review process is typically under, information should be presented in the most optimal configuration to support cognitive processes and the ability to retain and compare information to existing criteria or qualifications. In an attempt to quantify the differences in résumés visual layouts of information, number of saccades, total scanpath length, average saccade amplitude, and entropy will be compared across two résumés that vary in the spatial layout of the résumé review criteria. Therefore, the purpose of this study is to quantitatively compare how people review two different recommended styles of résumés by using saccade-derived eye tracking metrics. If cognitive processes can be better explained from these metrics, understanding which résumé layouts influence the reviewer's fit/no fit decision process

can lead to empirically-based résumé layout suggestions, which is presently lacking in the literature.

Hypotheses

Since there is little information about how résumé layouts affect saccade-derived metrics, it was unclear which résumé hierarchy would result in a more optimal scan pattern and layout when considering the fit/no-fit decision. Therefore, the following hypotheses were exploratory in nature.

Hypothesis 1: The vertically proximal résumé will have a lower number of saccades than the horizontally proximal résumés. Since the critical items of the vertically proximal résumé are all aligned on the same axis, it is predicted, "smoother" scanning will be observed and therefore lead to less saccades.

Hypothesis 2: The vertically proximal résumé will result in shorter scanpaths and average saccade amplitude than the horizontally proximal résumé. This is predicted since the vertically proximal résumé has résumé review criteria closer together than the horizontally proximal résumé.

Hypothesis 3: The vertically proximal résumé will have a smaller normalized transition entropy value than the horizontally proximal résumé. Normalized transition entropy will be calculated as previous evidence indicated transition entropy is affected by the item of interest's location [6]. Since the vertically proximal résumés has résumé review criteria closer together, it is predicted the randomness of the scanning will be less than its horizontally proximal counterpart.

METHOD

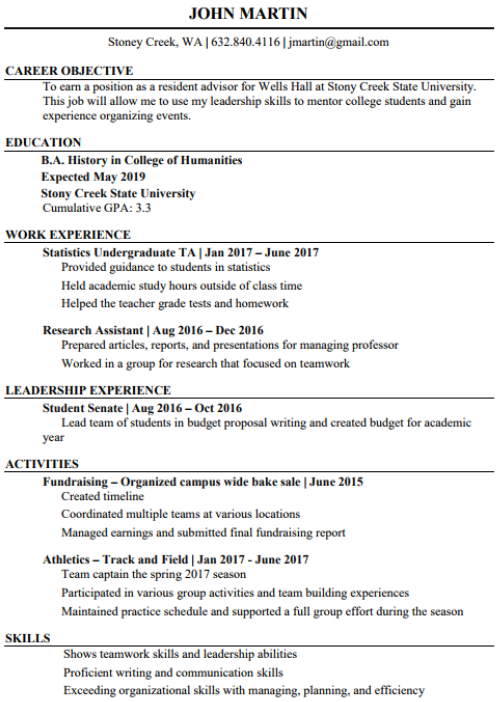
Participants

Twenty participants (13 male, age range of 20-24 years) participated in this study. All participants were either upper class undergraduates or graduate students at Clemson University. All had normal or corrected to normal vision. Participants were not compensated for their time as participation was completely voluntary.

Stimulus

Participants were tasked in making a fit/no fit decision and finding a specific résumé review criterion for two different, candidates applying to independent job openings. Both job openings were for a Resident Advisor position at two completely independent universities. Both universities were fictitious in order to avoid any familiarity bias. The résumés used to make this decision are in Figure 1 and Figure 2. Both job candidates had the same amount and type of work experience, but with different company names, dates, and action words. Résumés were created this way so to ensure that any differences in eye movements occurred due to the layouts, not the specific information.

The résumé in Figure 1 followed a vertical proximity visual hierarchy. In this case, the most important information, which includes current and previous position titles, start and end dates, and education [1], was placed serially down the left most side of the page, with specific details pertaining to each criteria being vertically spanning across the width of the résumé. The résumé in Figure 2 followed a horizontal proximity visual hierarchy as the most important information was distributed horizontally across the résumé with specific details filled in directly underneath it and spanning the entire width of the résumé.



JOHN MARTIN

Stoney Creek, WA | 632.840.4116 | jmartin@gmail.com

CAREER OBJECTIVE

To earn a position as a resident advisor for Wells Hall at Stony Creek State University. This job will allow me to use my leadership skills to mentor college students and gain experience organizing events.

EDUCATION

B.A. History in College of Humanities
Expected May 2019
Stony Creek State University
Cumulative GPA: 3.3

WORK EXPERIENCE

Statistics Undergraduate TA | Jan 2017 – June 2017
- Provided guidance to students in statistics
- Held academic study hours outside of class time
- Helped the teacher grade tests and homework

Research Assistant | Aug 2016 – Dec 2016
- Prepared articles, reports, and presentations for managing professor
- Worked in a group for research that focused on teamwork

LEADERSHIP EXPERIENCE

Student Senate | Aug 2016 – Oct 2016
- Lead team of students in budget proposal writing and created budget for academic year

ACTIVITIES

Fundraising – Organized campus wide bake sale | June 2015
- Created timeline
- Coordinated multiple teams at various locations
- Managed earnings and submitted final fundraising report

Athletics – Track and Field | Jan 2017 - June 2017
- Team captain the spring 2017 season
- Participated in various group activities and team building experiences
- Maintained practice schedule and supported a full group effort during the season

SKILLS

- Shows teamwork skills and leadership abilities
- Proficient writing and communication skills
- Exceeding organizational skills with managing, planning, and efficiency

Figure 1. The vertically proximal résumé.

JORDAN MAXWELL

North Franklin, OR 489.265.3278 jmaxwell@gmail.com

CAREER OBJECTIVE

To obtain a job as a resident advisor for James Hall at North Franklin State University. My leadership skills will help me mentor residents while I gain experience organizing events.

EDUCATION

North Franklin State University — Expected May 2018
B.A. History in College of Humanities — Cumulative GPA: 3.3

WORK EXPERIENCE

Research Assistant — Jan 2017 – June 2017
Completed presentations, reports, and articles
Assisted research administrator
Worked on team based research that focused on group collaboration

Algebra Undergraduate TA — Sep 2015 – Dec 2016
Provided assistance to students in algebra
Helped students during study hours outside of class time
Graded assignments and tests as the teacher requested

LEADERSHIP EXPERIENCE

Student Government — Jan 2016 – Mar 2016
Lead team of undergraduates
Worked on grant writing
Proposed allowance for academic year

ACTIVITIES

Fundraising – Organized community car wash — May 2016
Coordinated multiple groups at different stations
Established timelines
Handled revenue and completed the earnings report

Athletics – Cross Country Team — Aug 2016 – Nov 2016
Team captain during the 2016 season
Involved in multiple team building exercises and group activities
Maintained practice schedule
Ensured a full group effort throughout the season

SKILLS

Display leadership abilities and collaboration skills
Excel in both written communication and oral communication
Skilled in organization through productivity and project management

Figure 2. The horizontally proximal résumé.

Apparatus

Eye tracking metrics were collected by a Gazepoint GP3 eye tracker which is an unobtrusive desk mounted device that uses pupil/corneal reflection to track a person's eye movements as they reviewed the résumé. The eye tracker's refresh rate was 60 Hz meaning data was collected every 16 ms with 1 degree of visual accuracy. All résumés were presented in their portable document form (PDF) and displayed on a 22" Dell monitor with a screen resolution of 1680 x 1050. Calibration with the eye tracker was done for each résumé viewing (twice per participant) with a 9-point calibration method.

Experimental Design

The present study was a within subjects design, meaning participants viewed two résumés. Data collected during this study included the difference in résumé viewing time, the fit/no fit decision, the accuracy of recalling a specific piece of résumé review criteria, number of saccades, saccade amplitude, scanpath length, normalized transition entropy coefficient across the two résumés. Stimuli presentation (i.e. résumé hierarchy) was counterbalanced for each participant. Additionally, a debriefing survey collected participant's rank of importance of résumé review criteria, previous experience reviewing résumés, and a free response question asking how the participants completed the résumé review task.

Procedure

Participants were asked either in-person or via email to take part in a short eye tracking study. After gaining complete voluntary agreement from the participant, an experimental session date and time was arranged based on the participant's availability. After being greeted, participants were briefed about the details of the experiment. After consenting to participate in the study, participants were then informed they would be responsible for reviewing applicants' résumé's for current job openings. Their job was to decide if each applicant was a good fit for the job position, where good fit was defined as "able to meet or exceed all the requirements of the open position". Additionally, they were tasked with finding the start date of the candidate's first job to ensure that specific information was actually searched for during the task. The hypothetical job position was a resident advisor role at the applicants' respective university. Participants were informed they would complete this task a total of two times and that the applicants were not competing for the same job opening, therefore the fit/no fit decision was independent.

After successful calibration of the eye-tracker, each participant reviewed one of the two résumés. After viewing the résumé long enough to make a decision, the space bar was pressed and a white screen appeared. Then, the participants told the experimenter the start date of the candidate's first job and their choice for the fit/no fit decision. This process was completed again for the second résumé. Figure 3 shows the experimental setup.

The user then completed a debriefing survey and was reminded that all responses were voluntary. Then the participant was asked if they had any additional questions or concerns for the experimenter. After addressing any concerns, the participant was dismissed. The experimental session took about 15-25 minutes.

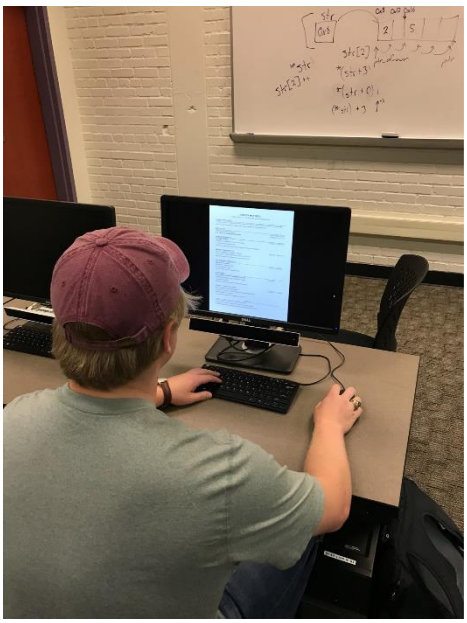


Figure 3. Experimental setup in which the participant views one of the résumés in randomized order.

RESULTS

All statistical analyses were performed using R 3.4.1 with a significance level of $\alpha = 0.05$. The measures of interest were time on task, accuracy, number of saccades, average saccade amplitude, total scanpath length, and transition entropy. Figure 4 and Figure 5 show how AOIs were determined for the vertically proximal and horizontally proximal résumé, respectively, and there were a total 11 AOIs per résumé.

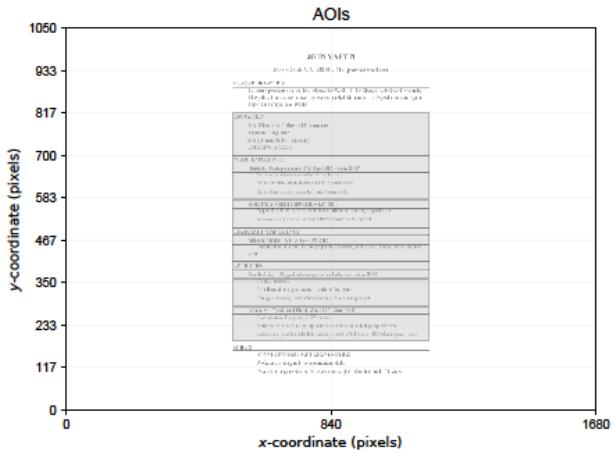


Figure 4. Pre-determined AOIs for the vertically proximal résumé.

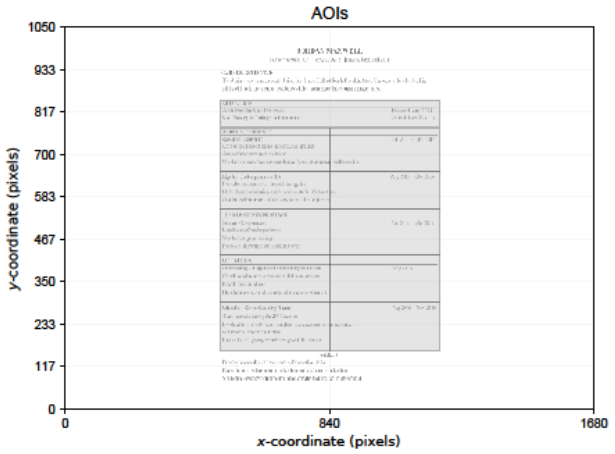


Figure 5. Pre-determined AOIs for the horizontally proximal résumé.

Fit/No Fit decision

Participants were asked to make the fit/no fit decision for each résumé they reviewed. Both résumés had an average fit decision rate of 90% meaning there was no significant difference ($t_{19} = 0$, $p = 1$; Fig. 6). The findings confirm the content of the two résumés was not significantly different and not a confounding variable on viewing pattern. Therefore, it was appropriate to move onto comparison analyses.

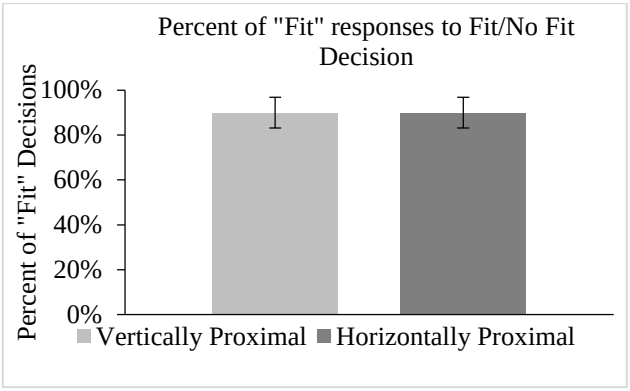


Figure 6. Comparison of the fit/no fit decision across both résumé types with respective standard error bars.

Résumé Review Time

Résumé review time was defined as the amount of time the participant spent reviewing the résumé. This data was automatically collected by the eye tracker and timing began and ended by the participant pressing the spacebar. The mean time on task for the vertically proximal résumé (M = 40.8 s) was higher than the mean time on task for the horizontally proximal résumé (M = 40.7 s) however the difference was not significant ($t_{19} = 0.04$, $p = 0.966$, n.s.; Fig. 7).

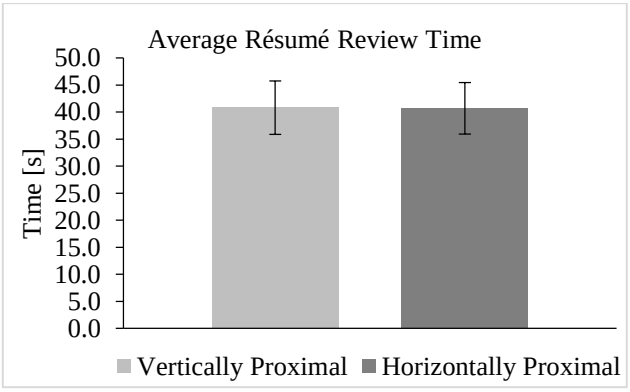


Figure 7. Comparison of time spent reviewing each résumé with respective standard error bars.

Accuracy of Recall Task

Participants were asked to identify and verbally recall the start date of the candidate's first job. The mean accuracy of the task across participants was calculated for the vertically proximal résumé (M = 65%) was lower than the mean accuracy for the horizontally proximal résumé (M = 70%). However, no significant differences between accuracy existed ($t_{19} = -0.397$, $p = 0.716$, n.s.; Fig. 8).

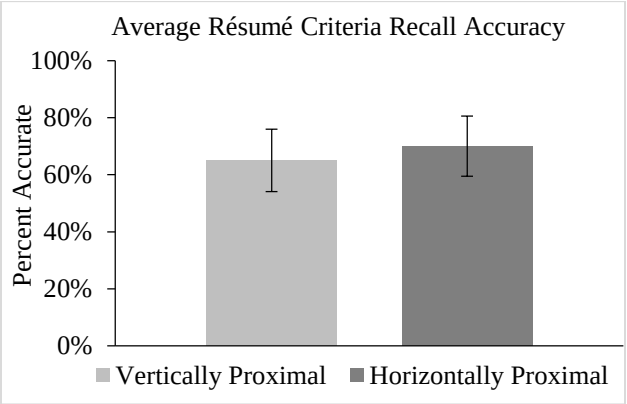


Figure 8. Comparison of average accuracy in recalling the start date of the candidates' first job position with respective standard error bars.

Number of Saccades

Number of saccades was calculated as the number of times eye movements that had a duration less than 30 ms as suggested by [12]. The vertically proximal résumé was predicted to have less saccades than the horizontally proximal résumé. The vertically proximal résumé had a larger mean number of saccades (M = 32) than the horizontally proximal résumé (M = 31). However, there was no significant difference between the two résumés (t_{19}= 0.32, p = 0.749, n.s.; Fig. 9).

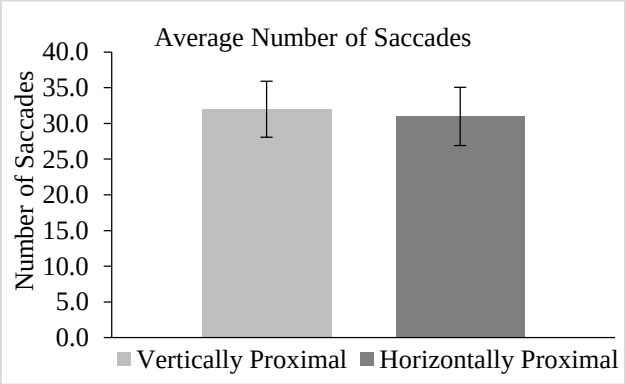


Figure 9. Comparison of number of saccades per résumé with respective standard error bars.

Average Saccade Amplitude

Average saccade amplitude was calculated as the arithmetic mean of all saccade lengths made when reviewing the résumé. It was predicted that, on average, the vertically proximal résumé would have shorter average saccade amplitude than the horizontally proximal résumé. The average saccade amplitude of the vertically proximal résumé (M = 143.615 pixels) was significantly smaller than the mean saccade amplitude of the horizontally proximal résumé (M = 232.408 pixels) (t_{19}= -8.62, p < .01; Fig. 10).

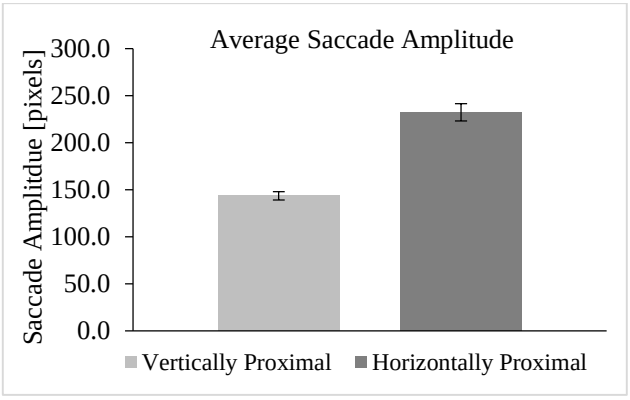


Figure 10. Comparison of average saccade amplitude for each résumé with respective standard error bars.

Total Scanpath Length

Total scanpath length was calculated as the total length of entire saccade-fixation-saccade sequence made while reviewing the résumé. It was predicted that, on average, the vertically proximal résumé would have shorter scanpath lengths than the horizontally proximal résumé. The average scanpath length of the vertically proximal résumé (M = 4679.7 pixels) was significantly shorter than the mean scanpath length of the horizontally proximal résumé (M = 7244.7 pixels) (t_{19} = -3.343, p < .01; Fig. 11).

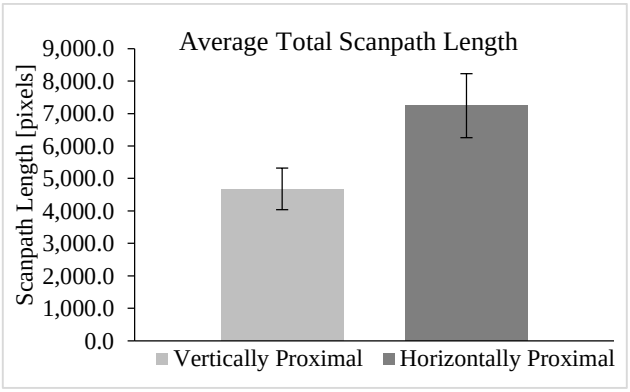


Figure 11. Comparison of the total scanpath length for each résumé with respective standard error bars.

Transition Entropy

Normalized transition entropy ($\widehat{H_t}$) was calculated per the suggestion of [7] as follows in order to compare across the two résumés [7]:

$$\widehat{H_t} = H_t \log_2(s)$$

Normalized transition entropy was used in order to compare values across résumés. Above, s is the number of AOIs and H_t is the transition entropy which is calculated as [8] suggests:

$$H_t = -\sum_{i=1}^{s} \pi_i \sum_{i=1}^{s} p_{ij} \log(p_{ij})$$

Again, s is the number of AOIs, π_i is the stationary probability of transitioning to AOI i over the long run and is estimated as

$$\pi_i = \frac{n_i}{n}$$

where n_i is the number of transitions to AOI i and n is the number of transitions total. Then, p_{ij} is the probability of transitioning from AOI i to AOI j. The vertically proximal résumé was predicted to have a smaller normalized transition entropy than the horizontally proximal résumé. The average normalized transition entropy for vertically proximal résumé (M = .37 bits/transition) was larger than the mean for the horizontally proximal résumé (M = .35 bits/transition), but there was no significant difference between the two résumés (t_{19}= 0.74, p = 0.468, n.s.; Fig. 12).

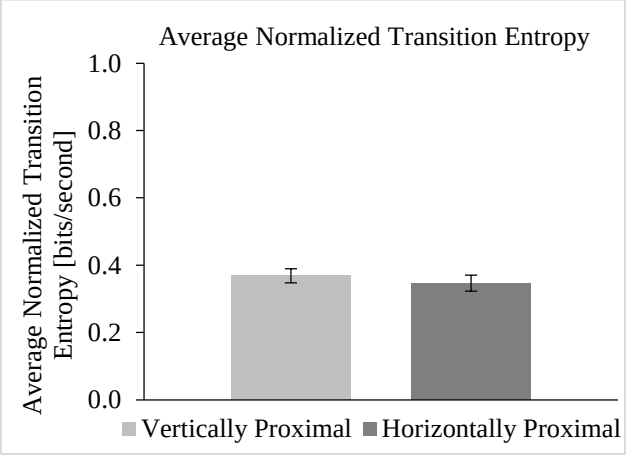


Figure 12. Comparison of normalized transition entropy between résumé types with respective standard error bars.

The maximum value of normalized transition entropy dependent on the number of AOIs and is calculated as:

$$Max(H_t) = \log(s)$$

For both résumés, this value was it is 1.041. This equation comes from calculating the transition entropy if all the transition probabilities followed a uniform distribution, meaning transitioning between AOIs was not Markovian in nature [7]. The minimum transition entropy value is 0 because if transitioning between AOIs was completely deterministic, there would be no randomness to quantify [7]. For further explanation, refer to [7, 8]. Normalized transition entropy values ranged from 0.21 to 0.53 from the vertically proximal résumé and 0.08 to 0.53 for the horizontally proximal résumé.

Debriefing Data

All participants completed a debriefing survey after reviewing the two résumés. Two participants indicated they had previous experiences reviewing résumés to make the fit/no fit decision. All but two participants indicated they used some sort of a review strategy to complete the task. Of those 18, twelve specifically cited they reviewed at least one of the résumé review criteria previously identified [3].

Analysis on Top Ranked Criteria

In the debriefing survey, each participant ranked the importance of each résumé review criteria when selecting a candidate as a good fit. In order to understand how the top ranked criteria influenced overall scan patterns, fixation duration and time until first fixation was founded for each participants' top ranked résumé review criterion as these metrics respectively indicate how long the reviewer spent processing and searching for the item [12]. These two metrics were compared between the two résumés. Data from one participant was excluded for this analyses because they added a criterion to the list and selected it as the most important. The total fixation duration of the top ranked item on the vertically proximal résumé (M = 5.22 s) was larger, on average, than the horizontally proximal résumé (M = 5.05 s) but it was not significantly different (t_{18}= -0.20, p = .841, n.s.; Fig. 13).

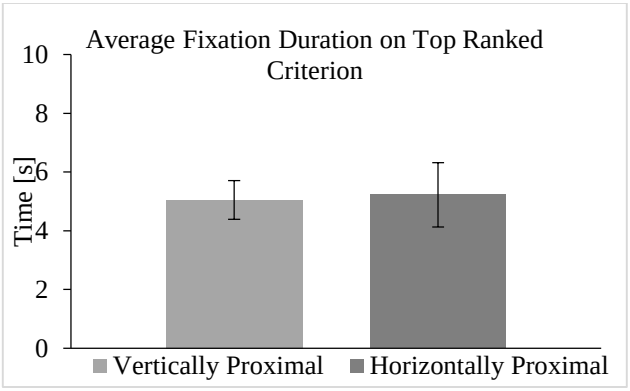


Figure 13. Comparison of average fixation duration on participants' top ranked criterion with respective standard error bars.

Time until first fixation on top ranked criterion for vertically proximal résumé (M = 4.31 s) was longer, on average, than for the horizontally proximal résumé (M = 1.65 s) but this difference was not significant at the significance level of α = 0.05 (t_{18}= 2.09, p = .051, n.s.; Fig. 14).

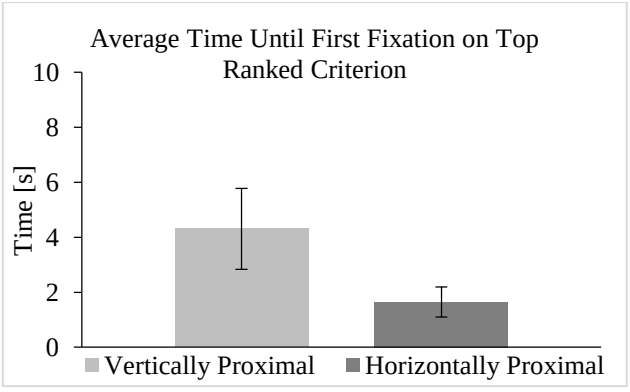


Figure 14. Comparison of average time until first fixation on participants' top ranked criterion.

DISCUSSION

The current study investigated how a visual layout of a résumé could influence the fit/no fit decision, length of résumé review, the résumé review criteria recall, and scan patterns. With respect to the first three topics, there was no significant difference between the two résumés. This could possibly be because participants had unlimited time to review the résumé , so making the fit/no fit decision and finding the start date of the candidate's first job could have been done more thoroughly than what has been observed with typical résumé reviewers [3]. Additionally, the content of the two résumés were similar by design, so the résumé layout was the only factor being tested.

When reviewing the eye tracking analysis, there was a significant difference in two eye tracking metrics: average saccade amplitude and total scanpath length. Average saccade amplitude, which has been found to indicate more meaningful or salient cues on the interface [12] and more pre-planning by the person [5], was significantly higher for the horizontally proximal résumé. This suggests the horizontally proximal résumé better distinguished résumé review criteria, making the distinguished items more salient. Additionally, these findings also indicated the horizontally proximal résumé may guide reviewers to pre-plan a set, systematic review strategy. Both of these implications therefore indicate résumé layout does influence the scan patterns of reviewers.

Total scanpath length has been found to measure scanning efficiency, with a higher value indicating a less efficient search, usually due to suboptimal layout [5, 11, 12]. This study found the horizontally proximal résumé to have a significantly larger average total scanpath length, indicating its layout may lead to a less efficient search. Considering résumé reviews are brief [3], this could lead résumé reviewers to review with less scrutiny and/or miss the résumé review criteria which would influence the fit/no fit decision. Although this was not the case in our experiment, the time also was not limited, so this large difference in scanpath could make a difference in fit/no fit decision and résumé review criteria recall if time was limited.

Although normalized transition entropy was not significantly different across résumés, their values was relatively low for both résumés, as based on the maximum possible value. This indicates that participants' transitions between key résumé review criteria, and therefore search in general, was not particularly random. This is expected as those who are scanning résumés tend to have an end-goal and would not be expected to just be viewing the résumé "freely" or in an exploratory manner [8]. Additionally, normalized transition entropy succinctly quantify an objective measure usually analyzed more holistically (e.g. reviewing eye tracking heat maps) and subjectively (e.g. debriefing data.) [3, 12, 13]. This finding helps inform empirically based résumé construction guidelines in practice. For example, if a professional wanted to compare résumé formats, he/she could have the résumés reviewed and compare transition entropy values. This finding also increases the application space of transition entropy in general.

As the figures show above, the standard error for average review time, number of saccades, fixation duration was large. This, along with the debriefing survey, indicates review strategies varied across participants. These findings support previous work that finds successful résumé construction depends on industry and reviewer due to variations in priorities so résumé selection rates may be improved if tailored to industry expectations [13]. This specific factor would need further investigation to understand its overall effect.

However, the fixation duration of the participant's top ranked item indicates participants did not change their internal review strategy and prioritization scheme between résumés. Also, these results also suggest that regardless of the résumé hierarchy, reviewers are only willing to spend a limited amount of time reviewing their top ranked criterion (average time on top ranked criterion accounted for about 13% of total résumé review time). Finally, the time until first fixation on a participant's top ranked criterion was approaching significance (p=0.051). The vertically proximal résumé was higher on average, which suggests résumé review criteria was more difficult to find than its horizontally proximal counterpart, even though it was spatially closer to other résumé review criteria.

Interestingly, these results suggest benefits for both of the résumé layouts in question. According to the results of this study, the vertically proximal résumé yielded results that suggest a more efficient search strategy, but with less of a pre-planned search strategy and potentially longer time to find top ranked résumé review criterion. At the same time, the horizontally proximal résumé yielded a potentially less efficient search, while also signifying participants pre-planned their search strategy which potentially accounted for the faster average time to top ranked résumé review criterion. These results indicate that résumé hierarchy affects how résumés are viewed. However, they do not necessarily suggest which résumé is best. Therefore, the present study was successful in guiding future research towards the correct topics of interest. Furthermore, they strengthen the argument that constructing a résumé should be done carefully and with regards to how the reviewer may search for and process information.

CONCLUSION

The aim of this work was to understand how résumé review was influenced by varying visual hierarchies. The findings were insightful, but not specific enough to deliver concrete suggestions when it comes to résumé construction. Future, and more specific work is needed to further understand this process and the factors influencing it as résumé review is a

commonality across all industries and business professionals. Empirical investigation in to this process is necessary in order to finally understand how résumé construction influences a candidate's chance at selection. It may also increase the rate of accurate "good fit" decisions, which could lead to less employee turnover and higher job satisfaction in the long run.

The evidence of this work suggests visual hierarchies on résumés can influence saccade-derived and scanpath based eye tracking metrics, which help explain search strategies. Although this work investigates a research question with a real need, generalizability of results and future work should be discussed.

Limitations

This initial eye tracking study evaluated concise metrics for résumé viewing patterns. In order to ensure that information organization was the difference of interest in this study, great care was taken to control the content of the two résumés. While creating résumés in this manner protected this study from confounding factors, the similar content in the résumés may not be as realistic as two randomly selected résumés from a real candidate pool. In the future, research should move towards understanding differences in scan path in a more naturalistic structure.

One of the main limitations in the present study is the amount of time each participant had to view each résumé. In a naturalistic setting, hiring managers spend as little as six seconds per résumé when making the initial fit/no fit decision. Participants were not limited on review time because (1) participants were students so résumé review experience was expected and confirmed to be limited and (2) understanding saccade-derived and scanpath metrics was on interest so the team had to ensure enough data was collected for each résumé for a sound analysis. Due to this, it is possible participants did not take the most efficient route when reviewing each résumé. In the future, researchers should limit the viewing time to understand the impact of a time limit.

Lastly, due to limitations in resources, the present study utilized undergraduate and graduate students as participants, not actual résumé reviewers. It could be possible that students employ very different search strategies and search for different résumé review criteria than professional résumé reviewers, even though our initial investigation showed there was a good amount of overlap. Specifically, the team feels résumé reviewers would complete the review process faster and at as a more automatized level of effort, meaning their searching may prove to be more efficient and dependent on factors like the visual hierarchy of the résumé . Future studies should reach out to this demographic to see if their patterns are more efficient and dependent on visual hierarchy.

Future Work

This work explored how résumé layouts affected a subset saccade-derived and scanpath metrics. Future work could include doing similar analysis on résumés that vary in different ways, like font style and size, indentation scheme, or density of content as these formatting structures have found to influence eye tracking metrics in previous work [2,10,11,13]. Additionally, since two of these metrics proved to be significantly affected by résumé layout, future work could look into other scanpath metrics like scanpath direction which determines a participant's directionality of search strategy (e.g. top-down vs bottom up) or saccade/fixation ratio as this evaluate the continuity of such a search [12]. More specifically, it may be of interest to investigate individual transition probabilities across participants and résumé structures. Transition entropy aggregates across all AOIs, so focusing in on specific transition probabilities between certain résumé review criteria, like those indicated to be the most important to a particular reviewer, would be of interest.

REFERENCES

1. Alan Baddeley and Graham Hitch. 1974. Working memory. Psychology of learning and motivation, 8 (Jan 1974), 47-89. DOI: http://dx.doi.org/10.1016/S0079-7421(08)60452-1
2. Soussan Djamasbi, Marisa Siegel and Tom Tullis. 2011. Visual hierarchy and viewing behavior: An eye tracking study. In *Human-Computer Interaction: Design and Development Approaches*, Orlando, FL, 331–340. DOI: http://dx.doi.org/10.1007/978-3-642-21602-2_36
3. Will Evans. 2012. Eye-tracking online metacognition: Cognitive complexity and recruiter decision making. (September 2015). Retrieved September 6th, 2017 from https://cdn.theladders.net/static/images/basicSite/pdfs/TheLadders-EyeTracking-StudyC2.pdf
4. Joseph H. Goldberg and Xerxes P. Kotval. 1999. Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, 24(1), (Jul 1998), 631–645.
5. Joseph H. Goldberg, Mark J. Stimson, Marion Lewenstein, Neil Scott and Anna M. Wichansky. 2002. Eye tracking in web search tasks: design implications. In *Proceedings of the 2002 symposium on eye tracking research & applications.* ACM Press, New Orleans, LA, 51-58. DOI: http://dx.doi.org/10.1145/507072.507082
6. Frank Keller. 2004. The entropy rate principle as a predictor of processing effort: An evaluation against eye-tracking data. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, 317–324.

7. Krzysztof Krejtz, Andrew T. Duchowski, Tomasz Szmidt, Izabela Krejtz, Fernando Gonzalez Perilli, Ana Pires, Anna Vilaro, Natalia Villalobos. 2015. Gaze Transition Entropy. *ACM Transactions of Applied Perception*. 13, 1 (Dec 2015). 1-20. DOI:http://dx.doi.org/10.1145/2834121
8. Krzysztof Krejtz, Tomasz Szmidt, Andrew T. Duchowski, and Izabela Krejtz. 2014. Entropy-based statistical analysis of eye movement transitions. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '14)*. ACM Press, Safety Harbor, FL, 158-166.
9. Jakob Nielsen. 2006. F-Shaped pattern for reading Web content. Jacob Nielsen's Alertbox (Apr 2006). Retrieved October 2nd, 2017 from https://www.useit.com/alertbox/reading_pattern.htm
10. Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology-Learning Memory and Cognition*, 26(3), 683-702.
11. Bing Pan, Helene A. Hembrooke, Geri K. Gay, Laura A. Granka, Matthew K. Feusner, and Jill K. Newman. 2004. The determinants of web page viewing behavior. In *Proceedings of the Eye tracking research & applications symposium on Eye tracking research & applications (ETRA '04)*. ACM Press, San Antonio, TX, 147–154. DOI: http://dx.doi.org/10.1145/968363.96839
12. Alex Poole and Linden Ball. 2006. Eye tracking in human-computer interaction and usability research: Current status and future prospects. *Encyclopedia of Human-Computer Interaction*, 211–219. http://dx.doi.org/10.4018/978-1-59140-562-7.ch034
13. Craig M. Ross and Sarah J. Young. 2005. Resume Preferences: Is it really "business as usual"? *Journal of Career Development* 32, 2 (Dec. 2005), 153-164. DOI: http://dx.doi.org/10.1177/0894845305279162
14. Dario D. Salvucci and John R. Anderson. 1998. Tracing eye movement protocols with cognitive process models. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*, Pittsburgh, PA, 923–928.
15. Frederick Shic, Katarzyna Chawarska, Jessica Bradshaw and Brian Scassellati. 2008. Autism, eye-tracking, entropy. *In 7th IEEE International Conference on Development and Learning (ICDL '08)*. Monterey, CA, 73–78.