

# **Visual Expertise is More Than Meets the Eye: An examination of holistic visual processing in radiologists and architects**

*Journal of Medical Imaging*

**Spencer Ivy,<sup>a\*</sup> Taren Rohovit,<sup>b</sup> Jeanine Stefanucci,<sup>c</sup> Dustin Stokes,<sup>a</sup> Megan Mills,<sup>d</sup> Trafton Drew<sup>c#</sup>**

<sup>a</sup>University of Utah, Department of Philosophy, Salt Lake City

<sup>b</sup>University of Oregon Department of Psychology

<sup>c</sup>University of Utah Department of Psychology

<sup>d</sup>University of Utah Department of Radiology & Imaging Sciences

## **Abstract**

### Purpose

One of the dominant behavioral markers of visual-expert search strategy, Holistic Visual Processing (HVP), suggests that experts process information from a larger region of space in conjunction with a more focused gaze pattern in order to improve search speed and accuracy. To date, extant literature suggests that visual search expertise is domain specific, including HVP and its associated behaviors.

### Approach

The current study is the first to use eye tracking to directly measure the HVP strategies of two expert groups, radiologists and architects, in comparison to one another and a novice control.

### Results

In doing so, we replicated and extended this prior research: visual expertise is domain specific. However, our eye tracking data indicate that contrary to this prior work, HVP strategies *are* transferable across domains. Yet, despite the transfer of HVP strategies, there is neither reduced search time nor greater accuracy in visual search outside of an expert's domain.

### Conclusions

Therefore, our data suggest that HVP behaviors are a particular form of visual search mechanism that, outside of an expert's native search-ecology, are not necessarily conducive to more general visual search success. It is in addition to explicit knowledge of an expert's domain, how to search, and where to search, that HVP strategies are their most effective for visual search success.

**Keywords:** Holistic Processing, Eye Tracking, Expertise, Visual Search, Gaze Contingent

\* First Author, E-mail: [spencer.ivy@utah.edu](mailto:spencer.ivy@utah.edu); [spencerivy93@gmail.com](mailto:spencerivy93@gmail.com)

# Last Author, Email: [trafton.drew@psych.utah.edu](mailto:trafton.drew@psych.utah.edu)

### *Acknowledgements*

We would like to thank the many radiologists at the University of Utah who participated in our study for their time and support. Additionally, we would like to thank the Salt Lake City based architecture firms, MHTN and FKRR, for their continued participation, time, and inspiration throughout the life of this project.

### *Declarations & Conflict Disclosure*

This research was in part funded by a Kickstart grant awarded to Trafton Drew. There are no conflicts of interest.

## 1 Introduction

There is no simple way to measure expertise. The challenge for researchers is that each expert is unique; each field of expertise affords its own unique requirements in skills, knowledge, and practice. But these challenges are not insurmountable. By isolating modestly limited features of skilled action, we stand to gain insight into the psychological mechanisms shared by experts generally. Accordingly, this study was concerned with *visual* expertise and the form of visual search behavior exhibited by two classes of visual experts. Specifically, we investigated *holistic visual processing* (HVP) as it relates to the search strategies of radiologists in comparison with architects and non-experts.

Holistic processing has long been established as a general behavioral marker for various forms of visual expertise. The visual expert is more sensitive to the global spatial array, and accordingly is often as or more sensitive to the whole object (of expertise) than its individual parts. This feature of expertise has been identified for faces, cars, fingerprints, elite sports, and among other kinds of expertise (Tanaka and Sengco 1997, Gauthier et al. 1998, Gauthier et al. 2002, Gauthier and Tarr 2002, Busey and Vanderkolk 2005, Mann et al. 2007).

One particular HVP theory of visual expertise posits that experts utilize two forms of parallel search representation to both see and search images differently than non-experts (Kundel et al., 2007; Gegenfurtner et al., 2011; Nodine & Mello Thoms, 2010; Sheridan and Reingold 2011). Experts engage in what is called a “global analysis” in which peripheral and parafoveal visual information delivers a holistic presentation of the visual search field. From the holistic perspective, targets and anomalies in the search field ‘pop-out,’ which causes a Gestalt resulting in the second form of search – “focal-feature analysis” wherein present and genuine anomalies are rapidly vetted and then selected (Palmer, 1990; Drew et al., 2013). The classic eye-tracking

measures indicating the presence of HVP are greater saccadic amplitudes, faster times to first fixation, and more efficient scan path ratios (Castelhano & Henderson, 2007; Gegenfurtner et al., 2011; Brams et al., 2019). These HVP-associated behaviors have been found in many kinds of visual experts including basketball players (Ryu et al., 2015), military personnel (Godwin et al., 2015), video-gamers (Marzouki, Dusayct, Chancaeux, & Mathot, 2017), chess players (Reingold, Charness, Pomplun, & Stampe, 2001), air traffic controllers (Van Meeuwen et al., 2014), and most prevalently in medical imaging sciences like radiology (Kundel et al., 2007; Sheridan & Reingold 2011; Donovan & Litchfield, 2013).

A common strategy for isolating the foregoing HVP behaviors is to use gaze-contingent viewing (GCV) windows that control for an expert's increased visual span. The purpose of GCV is to occlude the collection and use of parafoveal and peripheral visual information in order to stop holistic processing (Nodine, & Kundel, 1975; Viviani & Swensson, 1982; Brams et al., 2019). In GCV conditions, only a small window of visual information is presented to a viewer at the center of their point of fixation, occluding all other visual information. Consequently, by restricting collection of non-foveal visual information, the two parallel search strategies associated with HVP should be stymied because the global analysis is hindered, and the holistic pop out of targets is undermined. In effect, by imposing GCV on experts who utilize HVP, those experts' natural search strategies should be suppressed and begin to look much more like that of non-experts.

In our first study of this phenomenon (Ivy et al., 2021), we hoped to find evidence of HVP in a class of visual experts who had yet to be formally researched: architects. To test architects, we developed a novel search task controlling for architects' visual search strengths in geometric visualization, 3-dimensional design and parallel perspectives with what are called "Zero-Point

Perspective” images. \* As we expected, and is re-reported beside this study’s new results, architects performed significantly better and were far more sensitive to the presence of targets than a naïve group on the same task. However, we were surprised to find that architects exhibited significantly lower saccadic amplitudes, slower reaction times, and insignificantly different scan path ratios and times to first fixation compared to the naïve group. Thus, contrary to our initial expectations, architects exhibited an absence of HVP search behaviors compared to the naïve group.

The current study is a follow-up to these original findings focused on radiologists for whom the use of HVP search strategies is a well-established phenomenon (Sheridan & Reingold, 2011). Additionally, this study is the first to compare the domain-specific features of search behavior between two distinct expert groups of images within and between their respective domains. Our purpose in directly comparing groups was to examine the overlaps and inconsistencies of visual expertise across populations of visual experts and, by focusing on HVP, to understand how and when HVP behaviors are beneficial to visual search. A related question was whether the HVP strategy typically associated with radiologists’ medical image search would transfer to the perspective search. Understanding the interchangeability and transferability of specific visual search strategies would have a direct impact on practiced medical image search. For instance, if there were a shortage of medical readers of a specific domain, it would be helpful to know whether readers from an alternative domain could be recruited to effectively help the shortage. Further, investigations of the overlaps of visual search strategy across domains could illuminate general visual skills associated with search-success that may be worth implementing into training paradigms.

---

\* Zero-point perspective images are 3-dimensional representations of boxes held in space, all parallel with one another. Should the boxes be in motion, they would converge at the center of the image – the ‘horizon point.’ The specific design details of the task are discussed in greater detail below in the Method section.

We hypothesized that radiologists would not exhibit HVP strategies outside of their domain of expertise because previous work has suggested that visual expertise is non-transferable across domains (Nodine & Krupinski, 1998; Robson et al., 2021). Famously, Nodine and Krupinski (N&K) had radiologists search images for WALDO in Where's Waldo puzzles, and for the word 'NINA' artistically hidden in cartoon scenes. N&K found that for these generalized search tasks, "there was no statistically significant difference in detection performance between radiologists and laypeople" (p. 1).

However, where the WALDO and NINA images solicited general and non-expert search behaviors, the images of this current study were specific to the training and skills of our visual expert populations: radiologists and architects. That is, we collected data from two expert populations using images specific to the tasks and training those visual experts exhibit on a day-to-day basis as opposed to N&K's comparison of one expert group against a naïve population manipulated by WALDO/NINA. The idea behind this investigation was to determine whether, specific to images that elicit visual expertise, there would be a transferability of skill across domain.

Accordingly, across both sets of images where visual expertise *does* occur, would HVP search strategies continue to be recruited to the benefit of the expert viewers? Inspired by this question, we hypothesized, following N&K's original research that there would not be a transfer of search strategy; HVP would not appear in out-of-domain image searches. Thus, we sought to find whether *any* portion of expert visual search behavior would transfer across specifically expert domains. And, contrary to both N&K and our hypothesis, we found that there was indeed a transfer of search behavior but not a transfer of search expertise. The claim that visual expertise is domain

specific appears to be true, but strong claims should be tempered. Search behaviors transfer, but it takes more than mere behavior to make an expert within a domain.

## **2 Method**

### *2.1 Participants*

Data were collected from 78 participants with normal or corrected-to-normal vision. The sample size was based on power calculations derived through G\*Power (Version 3.1.9.2; Faul, Erdfelder, Lang, & Buchner, 2007). This calculation was done for F-tests with repeated measures that were within-subjects with 3 groups. An effect size of 0.5 and an alpha of 0.05 error resulted in groups of 24 with power at 0.80. Of the 78 participants, we collected data from two populations of visual experts (architects and radiologists) and one of naïve subjects (undergraduate students). Architects ( $n = 27$ , 7 female, 20 male) held either master's degree or a license to practice architecture. The data from the naïve and architect groups has been previously published in an original study (Ivy et al., 2021). The average age of architects was 42 ( $SD = 11.63$ ) who collectively averaged 19 years of experience ( $SD = 12.12$ ). Due to poor calibration of the eye-tracker, five of the 27 architects' data were not analyzed. Radiologists ( $n = 23$ , 11 female, 12 male) were either board certified to practice radiology, or had graduated medical school and were in residency at the University of Utah Medical School to acquire their certification. The average age of radiologists was 32 ( $SD = 4.4$ ) who collectively averaged 4 years of experience ( $SD = 3.36$ ). Undergraduates ( $n = 28$ , 20 female, 8 male) were recruited through the psychology department participant pool for course credit and had no experience in either medical imaging or architecture and collectively averaged age 19 ( $SD = 1.93$ ). Due to poor calibration of the eye-tracker, four of the undergraduates' data were not analyzed. None of the architects had any formal training in medical imaging; only one of the radiologists had an undergraduate degree in art, none had any

formal experience in architecture. Architects were compensated with \$30 for their participation, radiologists with \$50, and undergraduates received course credit.

## *2.2 Materials*

Data were collected using an Eyelink Plus 1000 (SR Research, Ontario Canada) eye tracker with a 1000Hz sampling rate and spatial resolution of 0.5 degrees of visual angle through two laptops. One laptop operated on the Eyelink OS (Version 5.15), and the other ran the experiment through MATLAB's (version R2017a) Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli 2007). A headrest was used to stabilize participants' heads and keep calibration consistent across trials. The images used in trials were of a consistent size: 28.3 cm x 28.3 cm, subtending a visual angle of  $24.6^\circ$  in either direction. All images were presented on a 21" monitor in 1920 x 1080 resolution without pixelation. Ambient room lights were set low with a single lamp in an otherwise dark room to simulate the lighting present in radiology reading rooms. On some trials, a gaze-contingent window was used to occlude vision of the monitor beyond a centered point of fixation. The nonoccluded area of the gaze-contingent window slowly faded to gray following a bivariate Gaussian circle whose half-width at half-maximum measured  $1.8^\circ$ . At  $1.8^\circ$  from the point of fixation, the gaze-contingent window was half gray and continued to fade until reaching 100% occlusion at  $4.9^\circ$ .

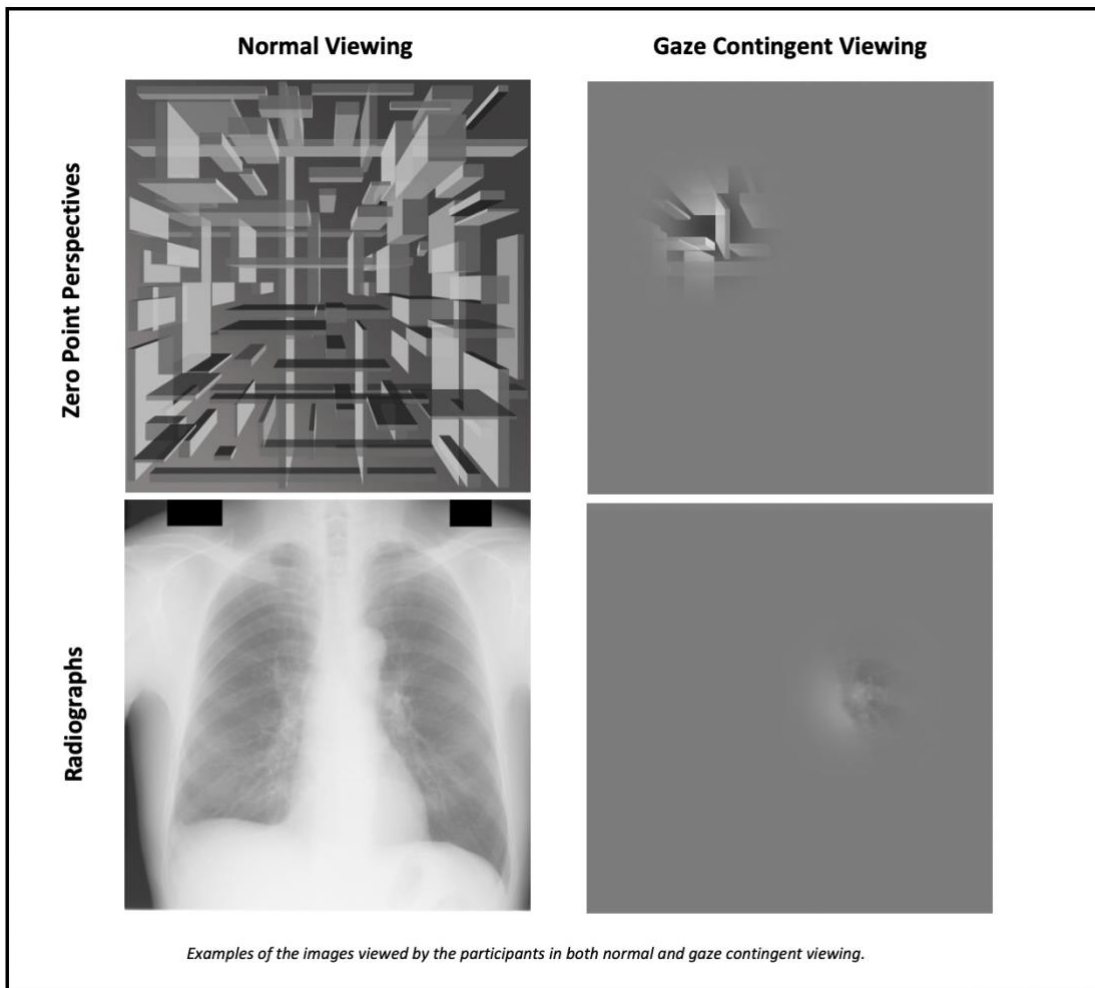
Twenty-two single view chest radiographs in the PA projection with a varying degree of search difficulty were drawn from the Japanese Society of Radiological Technology database (Shiraishi et al., 2000), of which 11 contained a malignant lung nodule. The difficulty of these images ranged based on the relative subtlety of targets established by Shiraishi et al. (2000) from 1 (most subtle/difficult) to 5 (least subtle/difficult). Participants were given an equal distribution of easier and more challenging images within each block. The average difficulty rating of images

within each block was 3, but always included images from both limits of the 1-5 range. These images were not calibrated to GSDF, but were both consistent with the high-resolution monitors widely used in much of the medical image perception literature. Although this is a limitation, we ultimately do not believe it is a crucial factor in the reported results as the focus of this study was not on overall radiologist performance, but examining how radiologists' performance change with our experimental manipulation of viewing condition, and how these changes relate to changes in another expert group (architects) to the same manipulation. We also have some evidence that our radiologist observers performed the task qualitatively similar to observers in the Japanese Radiology Survey (Shiraishi et al., 2000), who viewed these same cases on film. Specifically, this database separated cases based on nodule subtlety (1-5) and validated this categorization with 20 radiologist viewers. We used a mixture of cases from the intermediate subtlety categorization and found that performance was similarly very strongly modulated by this manipulation with our radiologist observers, but not with our architect observers.

Twenty-two grayscale perspective images were created in the 3DS Max software suite containing between 40 and 150 geometric shapes in zero- or one-point perspective with one another. All lines of these shapes were in parallel with one another and, should they have been in motion, would have converged at the horizon point at the center of the screen (see Figure 1, also see Ivy et al., 2021 for additional details). Eleven of these images contained a single geometric target out of perspective with the rest of the shapes. In choosing the zero-point perspective models to test for architects' visual expertise patterns and search behaviors, we worked with architects to construct a paradigm that would target both the *forms* of images that they examine and search on a daily basis as well as manipulate the *strategies* of visual expertise that they develop through experience and training. Namely, our zero-point search task highlighted visual skills associated



with identifying holistic congruence and spatial parallels within 3-dimensional digital designs. The design of this task was constructed to parallel architects' typical day to day search strategies and tasks. Architects frequently look at abstract forms in layers of blueprints, digitally designed objects, and AutoCAD models. Similar to the radiology images, perspective images were ranked on a scale of target subtlety and distributed analogously across the counterbalanced participant groups to include both challenging and easy search arrays. The perspective task was selected after discussions with these pilot participants who suggested that finding the out-of-perspective target was consistent with their on-the-job skills in reading blueprints and design.



**Fig. 1** Examples of image types that participants were tasked with searching for targets in their respective viewing conditions

### 2.3 Procedure

All procedures were reviewed and approved by the University of Utah Institutional Review Board and complied with the tenets of the Declaration of Helsinki. The experiment began by acquiring informed consent from each participant. Before beginning the experiment, each participant completed a 9-point calibration and validation procedure until the eye tracker was calibrated within a margin of  $.5^\circ$  visual angle. Participants were then told that they would search for targets in two types of images (radiographs and perspective images) and in two different viewing conditions (normal viewing and gaze-contingent). In the normal viewing condition, the entire screen was visible to the participant, whereas in the gaze-contingent condition, all participants could see was a small window focused at the center of their point of fixation (see Figure 1). In this gaze-contingent condition, participants could move the unoccluded window from point to point by moving their eyes to different parts of the screen. This was controlled by the eye-tracker which followed the participants' gaze and updated the image on the screen at 120Hz accordingly. For radiographs, participants were told to search for tumors or signs of cancer. For perspective images, they were told to search for the box that that was out of perspective with the rest. Upon finding a target, participants were told to click on it, then record their confidence from one to six in the answer they provided; one being the least confident, and six the most confident. Half of all trials had a target, and half did not. Participants were aware of the fact that there would only be a target present in about half of the trials. If participants could not find a target, they were instructed to click a box on the left side of the screen stating, "Nothing wrong? Click here," then give their confidence rating.

In order to acquaint participants with the types of targets that they would search for, participants were given 8 total practice trials – four prior to the perspective task, and four prior to

the radiograph task. A target was present in each practice trial and it was only in practice that participants were shown the correct location of the target. Once the practice was complete, each experimental block consisted of 22 trials of a single image type (perspective or radiograph). These trials were divided into 2 groups of 11 images: 11 in normal viewing and 11 in gaze contingent viewing. Once participants completed a 22-image block of either perspective or radiograph images, they would then perform the same task for the alternative image type. The ordering of groups within blocks was counterbalanced across participants and the images were randomized within each counterbalanced group. No participant was exposed to the same image across tasks or within blocks. After participants had completed searching all 44 images, the experiment was concluded with a single post-performance interview that aimed at gathering qualitative data on general confidence and individual strategies across tasks and conditions.

### **3 Results**

The presentation of results is organized by image type (perspective then radiograph), followed by a comparison of effects within each group across tasks. We calculated mixed-effect analyses of variance (ANOVAs) in two formats. First, for data comparing all three groups to each other (architects, radiologists, naïve), we report omnibus ANOVA results consisting of the three groups and two viewing conditions (GCV, normal). Second, to provide closer analyses of important effects, when comparing one group with another, we report mixed ANOVA results in factors of only two groups and two viewing conditions. Finally, when comparing the effect of interactions between viewing conditions and image types within each group (i.e., the differences between an observer group's performance on the radiology task vs. that same observer group's performance on the perspective task), we report the values of a paired t-test. Descriptive statistics for these data are presented in Tables 1 and 2. Although we initially computed both accuracy on

target present and absent trials and  $d'$ , we found that the qualitative pattern of results was identical of these metrics of accuracy. Therefore, to simplify, we report only  $d'$ , which is a probability measure weighing the likelihood of selecting a correct answer given a target's absence or presence, often used to measure sensitivity to perturbances in target search (Stanislaw & Todorov, 1999). Because search speed is a primary measure of HVP success, we also report reaction time as an additional measure to countenance the likelihood of HVP significantly amending search. Results concerning only the architect and naïve observer groups were originally reported in Ivy et al. (2021).

**Table 1** Perspective Images. Means and standard deviations in relation to observer group, view type, and target presence when viewing the perspective images

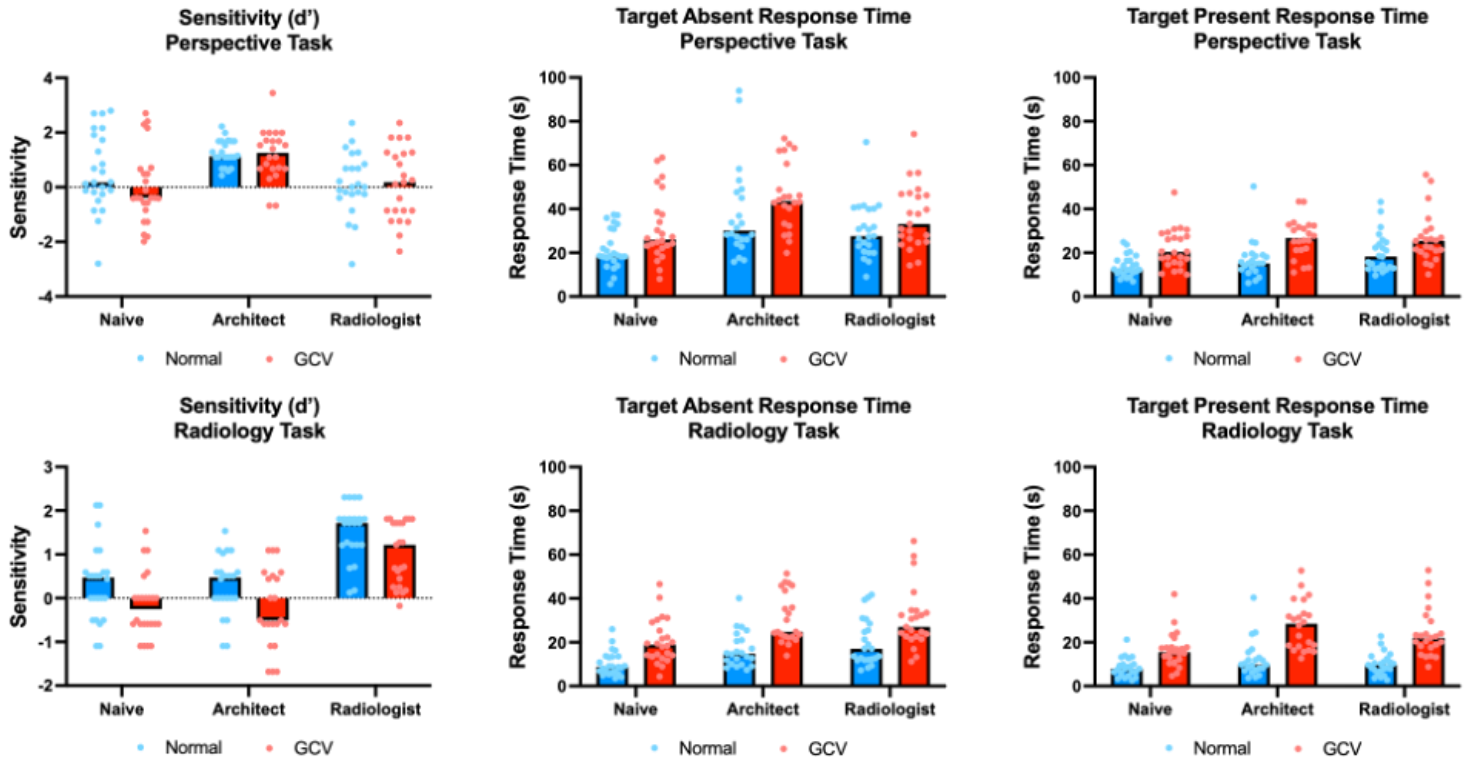
Perspectives	Architects				Radiologists				Naïve			
	Gaze-contingent viewing		Normal viewing		Gaze-contingent viewing		Normal viewing		Gaze Contingent viewing		Normal viewing	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>d'</i>	1.19	0.93	1.25	0.48	0.12	1.32	0.15	1.14	(0.02)	1.32	0.55	1.41
Saccadic amplitude (target present)	1.57	0.35	2.24	0.42	2.98	0.54	3.93	0.73	2.07	0.55	3.27	0.76
Saccadic amplitude (target absent)	1.63	0.39	2.35	0.58	2.94	0.54	4.17	0.96	2.25	0.47	3.51	0.77
Reaction time (s) (target present)	26.69	8.69	16.85	8.88	26.62	11.36	20.08	8.95	21.93	8.84	13.98	4.82
Reaction time (s) (target absent)	45.97	15.42	37.91	20.97	36.43	14.63	30.27	12.76	31.14	15.18	21.91	9.08
Time to first fixation	6.61	2.25	3.89	1.75	12.58	4.89	7.88	3.82	6.04	3.02	3.34	1.44
Decision time	14.69	4.44	11.05	5.98	16.89	8.99	13.73	7.85	14.27	7.50	9.15	4.55
Scan path ratio	14.98	7.43	12.99	5.31	16.14	8.38	15.92	8.47	15.20	10.98	12.84	7.01

Note. Averages for each variable were first computed withing participants, then averaged across participants. All values were rounded to the nearest hundredth

**Table 2** Radiograph Images. Means and standard deviations in relation to observer group, view type, and target presence when viewing the radiograph images

Radiographs	Architects				Radiologists				Naïve			
	Gaze-contingent viewing		Normal viewing		Gaze-contingent viewing		Normal viewing		Gaze Contingent viewing		Normal viewing	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>d'</i>	-0.24	0.89	0.29	0.69	1.03	0.71	1.51	0.63	-0.17	0.73	0.34	0.87
Saccadic amplitude (target present)	1.46	0.53	2.76	0.68	2.79	0.59	4.71	0.92	1.97	0.99	3.35	0.57
Saccadic amplitude (target absent)	1.52	0.38	2.69	0.75	2.86	0.60	4.75	0.95	2.14	0.52	3.89	0.82
Reaction time (s) (target present)	27.76	11.21	12.13	8.33	24.19	11.05	9.99	4.98	16.46	7.86	8.39	4.24
Reaction time (s) (target absent)	30.83	11.34	16.68	8.11	30.84	13.92	20.07	10.65	20.52	10.11	10.26	5.51
Time to first fixation	4.48	3.00	2.23	1.46	8.83	3.97	2.96	1.19	3.57	1.67	2.15	0.79
Decision time	15.64	7.26	7.01	6.07	16.13	11.17	6.38	4.06	9.87	5.70	4.38	2.92
Scan path ratio	6.10	2.46	4.95	4.10	10.34	6.89	4.91	3.92	5.73	3.51	6.22	4.81

Note. Averages for each variable were first computed withing participants, then averaged across participants. All values were rounded to the nearest hundredth



**Fig. 2** Average  $d'$  and reaction times for participants across task conditions and view types. Each dot represents a single participant's distinct performance of the respective trials.

### 3.1 Behavioral results for Perspective Images

With respect to  $d'$  and consistent with our hypothesis, there was a main effect of observer group ( $F(2,66) = 10.72, p < 0.01, \eta_G^2 = 0.15$ ) that appears to be driven by the fact that architects were significantly more sensitive to targets than both radiologists ( $p < 0.01$ ) and the naïve group ( $p < 0.01$ ). There was no difference between radiologists and the naïve group in  $d'$ , ( $p = 0.65$ ). Surprisingly, there was no effect of the GCV viewing condition on sensitivity ( $F(1,66) = 1.47, p = 0.23, \eta_G^2 = 0.01$ ), and the two factors (group, viewing condition) did not interact significantly (see Figure 2).

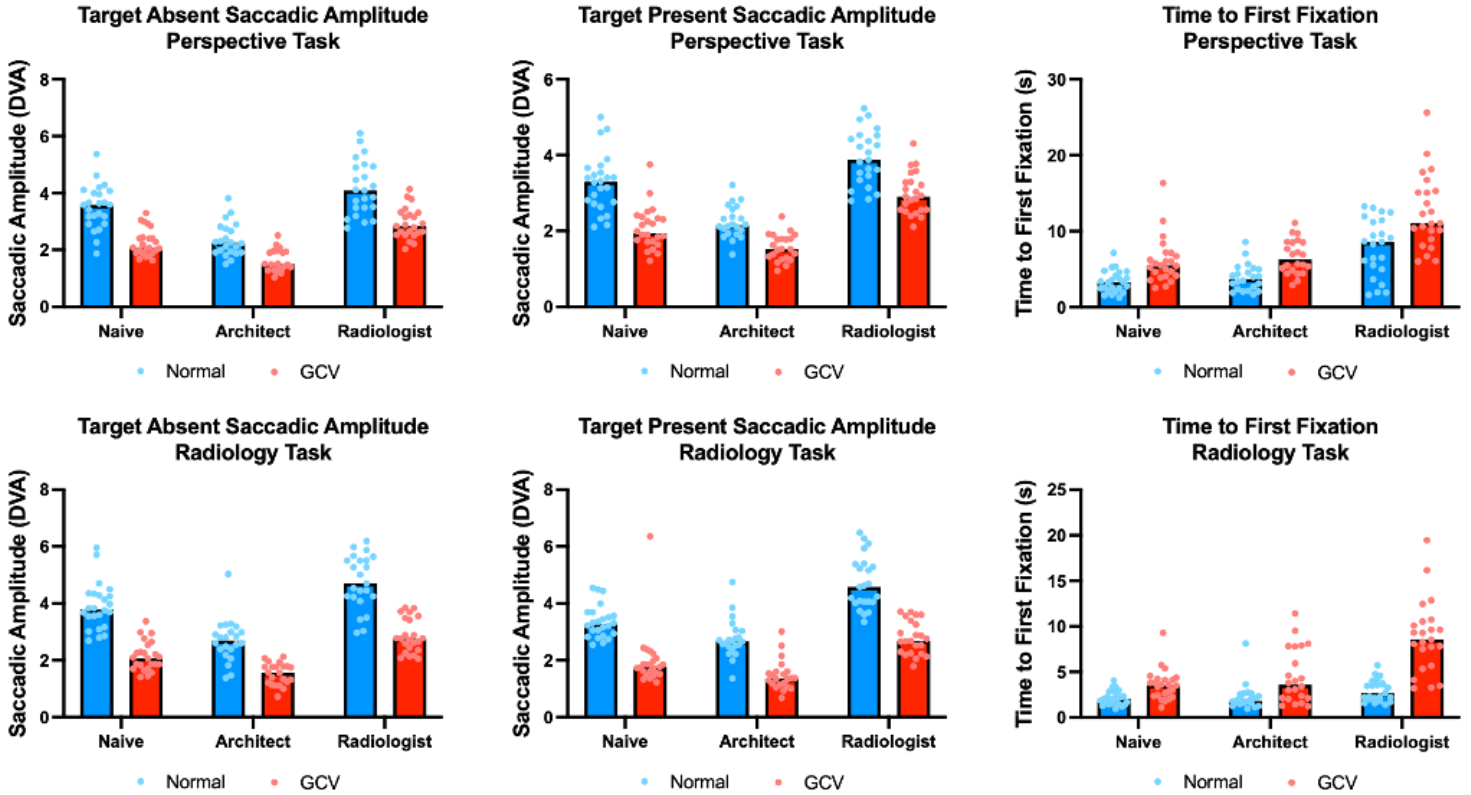
There were main effects of reaction time for both observer group ( $F(2,66) = 3.26, p = 0.04, \eta_G^2 = 0.07$ ) and viewing condition ( $F(1,66) = 52.34, p < 0.01, \eta_G^2 = 0.18$ ) in target present and target absent trials, but no effect on the interaction of observer group and viewing condition. In target

absent trials, architects were slower than the naïve group, ( $p < 0.01$ ). However, in target present conditions, no significant difference was found between the reaction time of architects and radiologists ( $p = 0.52$ ), nor between architects and the naïve group, ( $p = 0.60$ ), yet radiologists were slower than the naïve group ( $p = 0.02$ , see Figure 2).

### *3.2 Behavioral results for Radiograph Images*

There was a main effect of observer group ( $F(2,66) = 40.26, p < 0.01, \eta^2 = 0.37$ ), as well as of viewing condition ( $F(1,66) = 15.30, p < 0.01, \eta^2 = 0.10$ ) on  $d'$ , but the two factors did not interact significantly ( $p > 0.05$ ). These effects appear to be driven by radiologists: radiologists'  $d'$  was significantly greater than both the architects ( $p < 0.01$ ) and naïve group ( $p < 0.01$ ); and was more significantly affected by GCV than architects ( $p < 0.01$ ), and the naïve group ( $p < 0.01$ , see Figure 2).

With respect to reaction time, there was a main effect of viewing condition in target absent trials ( $F(1,66) = 132.12, p < 0.01, \eta^2 = 0.25$ ) and in target present trials ( $F(1,66) = 159.62, p < 0.01, \eta^2 = 0.37$ ). This is likely due to the fact that in both target present and absent trials, the naïve group was significantly faster than the radiologists ( $p < 0.01$ ) and architects ( $p < 0.01$ ). However, there was no significant difference in reaction times between radiologists and architects in both target absent trials, ( $p = .59$ ), and target present trials, ( $p = 0.24$ , see Figure 2). When comparing each observer group's performance across the perspective and radiology tasks, we found that there was a greater effect on the reaction time of architects in radiographs than in perspective images ( $p = 0.02$ ), and also for radiologists ( $p = 0.01$ ), but not for the naïve group ( $p = 0.94$ ).



**Fig. 1** Average Saccadic amplitude and time to first fixation for participants across task conditions and view types. Each dot represents a single participant’s performance of the respective trials.

### 3.3 Eye tracking results for Perspective Images

*Saccadic Amplitude* (SA) is a measure of the visual angle between movements of the eye and provides some of the strongest evidence of holistic visual processing: greater saccadic amplitude is associated with an increased ability to process information in the periphery (e.g., Gegenfurtner et al., 2011; Brams et al., 2019). For target present conditions, there was an effect on observer group ( $F(2,66) = 50.75, p < 0.01, \eta_G^2 = 0.55$ ), and viewing condition ( $F(1,66) = 212.37, p < 0.01, \eta_G^2 = 0.41$ ) as well as an interaction between observer group and viewing conditions ( $F(2,66) = 5.54, p < 0.01, \eta_G^2 = 0.03$ ) on SA. The same pattern was also identified in target absent trials (all  $p$ 's  $< 0.01$ ), so for simplicity we only report statistics for the target present trials. When searching in perspective images, the SA of architects was lower than both radiologists ( $p < 0.01$ )



and the naïve group ( $p < 0.01$ ). Architects were also less affected by GCV than radiologists, ( $p = 0.04$ ) and the naïve group, ( $p < 0.01$ ). The naïve group searched perspective images with larger SA than architects in both target absent trials ( $p < 0.01$ ) and target present trials ( $p < 0.01$ ). Interestingly, radiologists exhibited the largest SA when searching perspective images, significantly greater than the naïve group in target present trials ( $p < 0.01$ , see Figure 3).

*Time to First Fixation* (TTFF) is a measure of time spent between the beginning of a trial and the participant's first fixation upon a target (when present). There was a significant effect of viewing condition on TTFF ( $F(1,66) = 65.1, p < 0.01, \eta^2 = 0.24$ ), but no significant effect of observer group ( $F(2,66) = 0.83, p = 0.44, \eta^2 = 0.02$ ). The main effect of viewing condition appears to be driven by the fact that radiologists exhibited a slower TTFF than both architects ( $p < 0.01$ ) and the naïve group ( $p < 0.01$ , see Figure 3).

*Decision Time* (DT) is a measure of time between the participant's first fixation upon the target and their selection of a target in target present trials. In target present trials, there was a significant main effect of viewing condition ( $F(1,66) = 19.91, p < 0.01, \eta^2 = 0.10$ ). Between normal and gaze contingent viewing, architects were faster than radiologists ( $p = 0.01$ ), and the naïve were the fastest – quicker than both architects ( $p < 0.01$ ) and radiologists ( $p < 0.01$ ).

*Scan Path Ratio* (SPR) is a measure of efficiency for saccades in target present trials. It takes into account the average angle of each saccade and their distance from the eventual target with the expectation that, in GCV conditions search patterns will be more serial and in normal conditions far more efficient (Sheridan and Reingold, 2017). SPR was calculated by plotting points of search fixation, then averaging their sequentially relative distance to present targets over the course of search. There was no significant difference between any of the observer groups' SPR ( $F(2,66) = 2.12, p = 0.13, \eta^2 = 0.04$ ), nor was there a significant difference from viewing condition

on any of the three groups' SPR ( $F(1,66) = 1.11, p=0.29, \eta_G^2 = 0.01$ ). This was contrary to expectations for architects searching for targets within their domain of expertise.

### 3.4 Eye tracking results for Radiograph Images

SA in radiograph searches mirrored the results in perspective images in both target absent and present trials, for which we report the present trial statistics. There was a significant effect of observer group ( $F(2,66) = 43.27, p<0.01, \eta_G^2 = 0.47$ ), viewing condition ( $F(1,66) = 224.45, p<0.01, \eta_G^2 = 0.53$ ), and the interaction of observer group and viewing condition ( $F(2,66) = 3.54, p=0.03, \eta_G^2 = 0.03$ ). The interaction is likely driven by the fact that radiologists utilized greater SA than the architects ( $p<0.01$ ) and the naïve group ( $p<0.01$ ) across conditions. Moreover, the naïve group utilized greater SA than the architects in both target absent ( $p<0.01$ ) and target present trials ( $p<0.01$ ). The naïve group was also more affected by GCV than architects in target absent ( $p<0.01$ ), and target present trials, ( $p<0.01$ , see Figure 3). There was no difference in the effect of GCV on the naïve group across radiograph and perspective tasks,  $p=0.47$ . However, the SA of both architects, ( $p<0.01$ ), and radiologists ( $p<0.01$ ) was significantly more affected by GCV in the radiology task than in the perspective task.

*TTF*. For radiograph images there was a significant main effect for viewing condition ( $F(1,66) = 68.27, p<0.01, \eta_G^2 = 0.28$ ), but not for observer group ( $F(2,66) = 0.69, p<0.50, \eta_G^2 = 0.01$ ). Consistent with main effect of viewing condition, radiologists were more affected by GCV than the architects ( $p<0.01$ ), and the naïve group ( $p<0.01$ ). However, there was a significant interaction for TTF between observer group and viewing condition for radiologists compared to the naïve group ( $p=0.02$ ), bringing the radiologists' TTF very close to that of the architects and naïve participants in normal viewing conditions (see Figure 3).

*DT* in radiographs was similar to *TTF* in that there was a significant main effect of viewing condition across the three observer groups ( $F(1,66) = 19.9, p < 0.01, \eta_G^2 = 0.01$ ). Interestingly, the architects and radiologists *DT* was remarkably similar across conditions. Additionally, the naïve group was associated with a lower *DT* than radiologists ( $p < 0.01$ ) and architects ( $p < 0.01$ ). However, as expected given holistic processing theory, compared to the naïve group radiologists were significantly slower ( $p < 0.01$ ), more affected by *GCV* ( $p < 0.01$ ), and the two factors interacted significantly ( $p = 0.01$ ). Finally, the *DT* of architects, ( $p = 0.02$ ), and radiologists, ( $p < 0.01$ ), was significantly more affected by *GCV* in radiographs than in perspective images.

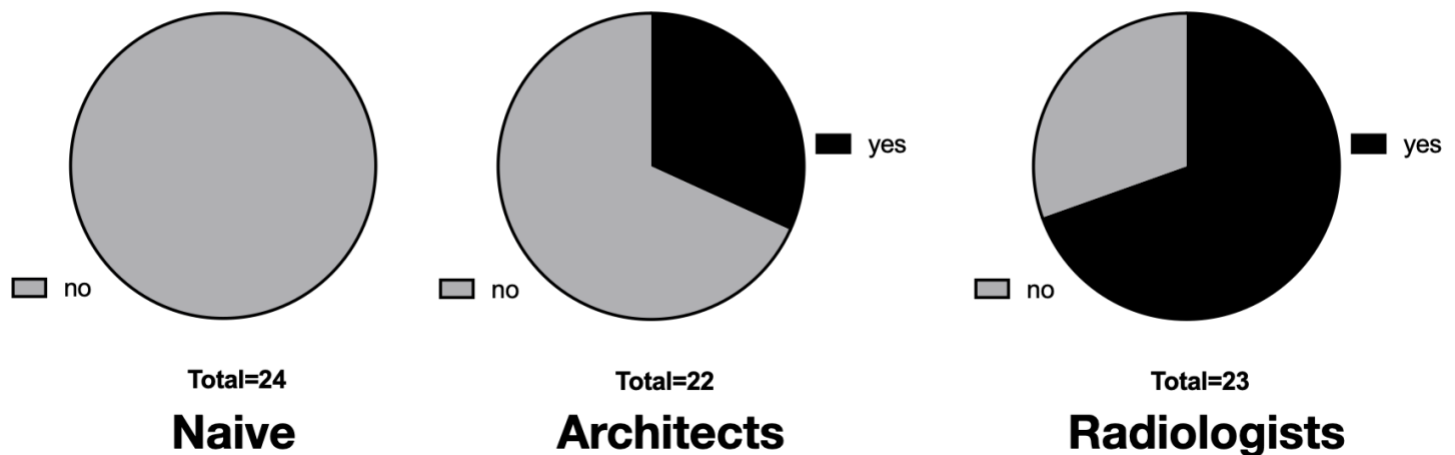
*SPR*. There was a significant main effect of viewing condition on *SPR* when viewing radiographs ( $F(1,66) = 7.54, p = 0.01, \eta_G^2 = 0.05$ ), but no effect of observer group. However, the two factors interacted significantly ( $F(2,66) = 6.07, p < 0.01, \eta_G^2 = 0.08$ ). This is likely due to the fact that radiologists were significantly more affected by viewing conditions than architects ( $p = 0.02$ ) and the naïve group ( $p < 0.01$ ).

### *3.5 Qualitative Reports*

When data collection was complete, we verbally interviewed participants to gauge their confidence across viewing conditions, strategies for each task, and the effect of *GCV* on those strategies. When asked the question, “Did you employ any strategies searching for targets in perspective images,” seven out of 22 architects used language strongly indicative of the phenomenology of holistic visual processing. These phrases included, “perform the gestalt,” “I zoomed my eyes out on the whole image to find the target,” “I just looked at the whole thing, then the target popped out to me,” etc. None of the naïve group used any language indicating a phenomenology consistent with holistic visual processing. A chi-squared test shows that these responses were significantly

different for the two groups, ( $\chi^2 (1, N = 46) = 4.74, p=.03$ ). A further analysis was done on the subset of architects who reported these strategies, but it did not yield any significant difference in performance between these architects and their counterparts. Despite using language indicative of holistic processing phenomenology, we found no evidence of architects having done so. In contrast, radiologists both used language strongly indicative of holistic processing, and exhibited the associated behaviors. Sixteen of 23 radiologists used similar holistic processing language in describing their search strategies when compared to novices, showing a significant difference in a chi-squared analysis, ( $\chi^2 (1, N = 46) = 9.41, p<.01$ , see **Figure 4**).

## Participants who expressed strategies consistent with holistic visual processing



**Fig. 2** Each diagram represents results from the chi-squared analysis performed on participants’ post-experiment reports. Black represents the portion of participants within each group who used language strongly indicative of a phenomenology consistent with holistic visual processing. Light grey represents the portion of participants who did not use the foregoing language.

## 4 Discussion

Our prior work suggested that HVP, as a marker of visual expertise, was not present in the behaviors of architect subjects (Ivy et al., 2021). Yet, lacking further data, we were unable to determine why this was the case. Because architects were more sensitive to the presence of targets,

and more accurate than the naïve group, we inferred that our perspective task did successfully capture architects' visual expertise. However, because the naïve group exhibited significantly greater saccadic amplitudes than architects and no difference in decision times or first fixations, our data suggest that HVP behaviors did not account for the architects' greater degree of accuracy on the task. This distinguished architects as a class of visual expert unlike those previously studied, a class that exhibits visual expertise without the use of HVP behaviors.

To explain this finding in our original dataset, we proposed two possible conclusions. First, it could be that architects would have used HVP, but something about the visual representation of the perspective task itself prevented it. Second, it could be that HVP behaviors in visual search are domain specific; some visual experts rely on greater visual span and holistic gestalts, while others – like architects (perhaps) – do not. The data of the present study, specifically those collected from radiologists, lead us to draw the second conclusion with an additional lesson: HVP strategies are domain specific, and their application across domains is not necessarily conducive to success.

This study's data showed that radiologists exhibited greater degrees of saccadic amplitude than both architects and the naïve group in the radiograph and perspective searches. Despite utilizing HVP behaviors in perspective images, radiologists were no more accurate nor more sensitive to targets than the naïve control. Accordingly, through our eye-tracking data, we can see that radiologists do apply some of their HVP strategies to search tasks outside of their domain, but the application does not improve their search ability. This finding rules out the possibility that the perspective images were exclusive of HVP. Radiologists were able to utilize HVP behaviors when searching the perspective images, but were not any better off than novices. This leads us to conclude that the usefulness of HVP behaviors is domain specific; architects do not use HVP behaviors and are better off for it within their own domain, and the inverse is true of radiologists

– they do utilize HVP and are better off for it, but only within their own domain. So, with respect to visual expertise, it appears that behavioral strategies like HVP are transferable, and yet, the strategies do not translate into success within alternative domains where expertise is lacking.

Most theories of visual expertise amount to the inclusion of three important strategies for success: working memory, information-reduction, and holistic processing (see Brams et al., 2019 for a helpful meta-analytic review). In other words, visual expertise consists in knowing where targets are (long-term working memory), knowing where targets will not be (information reduction), and how experts look for targets (holistic processing, Ericsson & Kintsch 1995; Haider & Frensch 1999; Kundel et al., 2007). The data of this study support the idea that the combination of these three strategies in visual search success is highly domain dependent, and in fact, the HVP portion is domain specific. Architects do not need HVP strategies to remain successful in visual search tasks associated with their form of visual expertise. Likewise, radiologists who do rely on HVP to remain efficient and successful in their own domain, apply their HVP strategies in out-of-domain searches but are not any better off for it.

These data lead us to believe that the domain specificities of visual expertise are as much a product of the habituated saccadic patterns in amplitude and search efficiency as they are in knowing how to search for targets, perturbances, and aberrancies. Radiologists know where in the lung structure they will find what they expect to find, and they know what those targets should look like. However, although the radiologists brought all of their best HVP search strategies to bear in finding perspective targets, they were no more accurate or efficient than the naïve group. Thus, visual expertise is more than meets the eye – or, at least, what makes a visual expert is not merely to be measured in how the eyes move within a visual search. The internal cognitive

processing of visual data and an expert's specialized domain knowledge are likely just as important as learned mechanical search patterns within the ecology of visual expertise.

The data of this study are a confirmation of the foregoing insight and are also, we hope, an inspiration for future research. Studies concerning the symbiotic relationship of mechanical and intelligent strategies in visual expertise stand to benefit from paradigms parallel to our own which include both behavioral and eye-tracking data collection. Expertise involves what experts know about targets in their domain, how they are trained to allocate their attention, and how efficiently they eventually become in utilizing those lessons as they collect visual data (Brams et al., 2019). At the intersection and interaction of these aspects of visual expertise there is still much left to be discovered.

#### *4.1 Limitations*

There are three possible limits to our data: sample size, age, and experience. With respect to sample size, our populations were twice the average of other studies of similar design. While the effects reported by similar studies on visual expertise have been drawn into doubt due to small sample sizes between 6-10 experts, our study of 45 total experts is, at least, less prone to this concern (Gegenfurtner et al., 2011; Brams et al., 2019). Further, it is unlikely that the average age difference between the naïve, architect, and radiologist groups had an effect on our data. Although visual ability does decline with age (Anderson, 2012; Owsley, 2011), the decline does not happen until well after the average age of our study participants, and is often on a different time frame than the decline of visual expertise. The decline of general visual acuity begins around age 60, whereas our architects were 42 years old on average, radiologists were 32, and the naïve group was 19 (McKendrick, Weymouth, & Battista, 2013). Accordingly, sample size and age are not likely to have affected our data.

The greatest limitation of our dataset is likely the difference in years of experience between the radiologists and architects. It is widely agreed that a respectable standard for ‘expertise’ is to have spent at least 10 years, or 10,000 hours within a domain with a constant desire to improve one’s abilities (Bereiter & Scardamalia, 1993; Yarrow, Brown, & Krakauer, 2009; Montero 2016; Ericsson, 2018). Our architect population far exceeded this standard with an average of 19 years of experience between them. The radiologists, however, averaged only four years of experience altogether. When meaning to compare apples to apples, this difference across this spectrum of expertise is concerning. However, our radiologist population was well within the average range between two and four years of experience for similar studies on the HVP strategies of radiologists (Krupinski, 1996; Kundel et al., 2007; Bertram et al., 2016). Further, the data from our radiologists is consistent with that of previous studies and therefore does not give us reason to think that the radiologists’ HVP strategies were affected by lack of experience (Drew et al., 2013; Sheridan & Reingold, 2011). Accordingly, an interesting direction for future research would be to measure the search strategies of intermediate architects as a contrast class to both their own expert counterparts and to the radiologists.

## **5 Conclusion**

What we set out to find in this study was the effect of visual expertise on searches within and without radiologists’ domain of practice compared to an alternative expert group: architects. We hypothesized that radiologists would use HVP strategies when searching for signs of cancer in radiographs, but not in the alternative task. Our data show that although radiologists utilized HVP strategies both in and out of their domain of expertise, those strategies only aided visual search success within their domain. From these data it is plausible to conclude that the visual search patterns of experts, like architects and radiologists, are domain and training specific. Specific search behaviors are likely developed as the result of training to find the different sorts of



targets that experts look for (i.e., what searches within their domain solicit), as well as what the visual field of those searches typically consists in (e.g., building blueprints or x-rayed lungs). Moreover, we found that HVP is applied across domain, but is not necessarily conducive to success. This finding is an invitation for new research: are any visual search strategies successfully transferable across domains of expertise, or is the success of each instance of search behavior ineluctably tied to a specialized domain? We hope that this study will be an inspiration to other researchers investigating this question and others like it in the spirit of uncovering the curious relationship between vision and expertise.

#### **List of Abbreviations**

*HVP*: Holistic Visual Processing  
*GCV*: Gaze Contingent Viewing  
*N&K*: Nodine and Krupinski  
*SA*: Saccadic Amplitude  
*DT*: Decision Time  
*TFFF*: Time to First Fixation  
*SPR*: Scan Path Ratio

## References

1. Tanaka, J.W. and Sengco, J.A. (1997). Features and their configuration in face recognition. *Memory and Cognition* 25, 583–592
2. Gauthier, I., Williams, P., Tarr, M.J., Tanaka, J. (1998) Training ‘greeble’ experts: a framework for studying expert object recognition processes. *Vision Research* 38, 2401–2428.
3. Gauthier, I. and Tarr, M.J. (2002) Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *J. Exp. Psychol. Hum. Percept. Perform* 28, 431–446.
4. Gauthier, I., Curran, T., Curby, K.M., Collins, D. (2003) Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience* 6, 428–432
5. Busey, T.A. and Vanderkolk, J.R. (2005) Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research* 45, 431–448
6. Mann, D.T.Y., Williams, A.M., Ward, P., Janelle, C.M. (2007). Perceptual-Cognitive expertise in Sport: A Meta-Analysis. *Journal of Sport and Exercise Psychology* 29, 457–478.
7. Bukach, C.M., Gauthier, I., Tarr, M.J., 2006. Beyond Faces and Modularity: The Power of an expertise Framework. *Trends in Cognitive Sciences* 10, 159– 166
8. Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic Component of Image Perception in Mammogram Interpretation: Gaze-tracking study. *Radiology*, 242, 396–402.
9. Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise Differences in the Comprehension of Visualizations: A Meta-Analysis of Eye-Tracking Research in Professional Domains. *Educational Psychology Review*, 23(4)523-552.
10. Nodine, C., & Mello-Thoms, C. (2000). The Nature of Expertise in Radiology. In *Handbook of Medical Imaging* (Vol. 1, pp. 859–894). SPIE Press.
11. Sheridan, H., & Reingold, E. M. (2017). The Holistic Processing Account of Visual Expertise in Medical Image Perception: A Review. *Frontiers in Psychology*, 8 :1620
12. Palmer, S. E. (1990). Modern theories of Gestalt perception. *Mind & Language*, 5(4), 289–323.

13. Drew, T., Evans, K., Võ, M. L.-H., Jacobson, F. L., & Wolfe, J. M. (2013). Informatics in Radiology: What Can You See in a Single Glance and How Might This Guide Visual Search in Medical Images? *RadioGraphics*, 33(1), 263–274.
14. Castelhana, M. S., and Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *J. Exp. Psychol.* 33, 753–763.
15. Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Williams, A. M., & Helsen, W. F. (2019). The Relationship Between Gaze Behavior, Expertise, and Performance: A Systematic Review. *Psychological Bulletin*, 145(10), 980–1027.
16. Ryu, D., Abernethy, B., Mann, D. L., & Poolton, J. M. (2015). The contributions of central and peripheral vision to expertise in basketball: How blur helps to provide a clearer picture. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 167–185.
17. Godwin, H. J., Liversedge, S. P., Kirkby, J. A., Boardman, M., Cornes, K., & Donnelly, N. (2015). The influence of experience upon information sampling and decision-making behaviour during risk assessment in military personnel. *Visual Cognition*, 23, 415–431.
18. Marzouki, Y., Dusaucy, V., Chanceaux, M., & Math t, S. (2017). The World (of Warcraft) through the eyes of an expert. *PeerJ*, 5, e3783.
19. Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, 12, 48–55.
20. van Meeuwen, L. W., Jarodzka, H., Brand-Gruwel, S., Kirschner, P. A., de Bock, J., & van Merrienboer, J. J. G. (2014). Identification of effective visual problem solving strategies in a complex visual domain. *Learning and Instruction*, 32, 10–21.
21. Donovan, T., & Litchfield, D. (2013). Looking for cancer: Expertise related differences in searching and decision making. *Applied Cognitive Psychology*, 27, 43–49.
22. Kundel, H. L., & Nodine, C. F. (1975). Interpreting Chest Radiographs without Visual Search. *Radiology*, 116(3), 527–532.

23. Viviani, P., & Swensson, R. G. (1982). Saccadic eye movements to peripherally discriminated visual targets. *Journal of Experimental Psychology: Human Perception and Performance*, 8(1), 113–126.
24. Ivy, S., Rohovit, T., Lavelle, M. *et al.* (2021). Through the eyes of the expert: Evaluating holistic processing in architects through gaze-contingent viewing. *Psychon Bull Rev* 28, 870–878.
25. Nodine CF, Krupinski EA. Perceptual skill, radiology expertise, and visual test performance with NINA and WALDO. *Acad Radiol.* 1998;5(9):603-12.
26. Robson SG, Tangen JM, Searston RA. The effect of expertise, target usefulness and image structure on visual search. *Cogn Res Princ Implic.* 2021 Mar 12;6(1):16.
27. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods.* 39, 175–191 (2007).
28. Brainard, D. H. The psychophysics toolbox. *Spatial Vision.* 10, 433–436 (1997).
29. Pelli, D. G. (1997). The VideoToolbox software for visual psycho-physics. *Spatial Vision*,10, 437–442.
30. Kleiner, M., Brainard, D., & Pelli, D. (2007). What’s new in Psychtoolbox-3? *Perception*, 36(ECVP Abstract Suppl), 14.
31. J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, et al., "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules", *AJR American Journal of Roentgenology*, 174(1), 71-74, (2000).
32. Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments & Computers*, 31(1), 137–149.
33. Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211–245.

34. Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 172–190.
35. Andersen, G. J. (2012). Aging and vision: Changes in Function and Performance From Optics to Perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(3), 403-410.
36. Owsley, C. (2011). Aging and Vision. *Vision Research*, 51(13), 1610-1622.
37. Mckendrick, A. M., Weymouth, A. E., & Battista, J. (2013). Visual Form Perception from Age 20 through 80 Years. *Investigative Ophthalmology & Visual Science*, 54(3), 1730.
38. Bereiter, C., & Scardamalia, M. (1993). *Surpassing Ourselves: An Inquiry Into the Nature and Implications of Expertise*. Chicago: Open Court.
39. Yarrow, K., Brown, P., & Krakauer, J. W. (2009). Inside the brain of an elite athlete: the neural processes that support high achievement in sports. *Nature Reviews Neuroscience*, 10(8), 585 596.
40. Montero, B. (2016). *Thought in Action: Expertise and the Conscious Mind*. Oxford University Press.
41. Ericsson, A. (2018). The Differential Influence of Experience, Practice, and Deliberate Practice on the Development of Superior Individual Performance of Experts. In *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge: Cambridge University Press.
42. Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Acad. Radiol.* 3, 137–144.
43. Bertram, R., Kaakinen, J., Bensch, F., Helle, L., Lantto, E., Niemi, P., & Lundbom, N. (2016). Eye movements of radiologists reflect expertise in CT study interpretation: A potential tool to measure resident development. *Radiology*, 281, 805– 815.