

Univerzita Karlova
Filozofická fakulta

Habilitační práce

Rozdělená pozornost a oční pohyby

2021

Mgr. Jiří Lukavský, Ph.D.

Obsah

Předmluva	4
1 Úvod	5
1.1 Odborný zájem	5
1.2 Zaměření jednotlivých metodologických studií	6
2 Teoretický kontext	7
2.1 Pozornost	7
2.2 Sledování více objektů	8
2.3 Oční pohyby při MOT	12
2.4 Modelování očních pohybů	13
3 Komentáře k prezentovaným studiím	16
3.1 Opakování očních pohybů	17
3.2 Statistické testování podobnosti očních pohybů	20
3.3 Oční pohyby a předvídaní pohybu	22
3.4 Pozornost a vliv distraktorů	24
4 Shrnutí metodologického přínosu	29
5 Aplikované analýzy očních pohybů	31
5.1 Vliv podoby studijních materiálů na efektivitu učení	31
5.2 Směr pohledu portrétovaných osob	33
5.3 Vnímání fotografií v kontextu	34
5.4 Vliv expertizy na zrakové vnímání	35
5.5 Shrnutí potenciálu metod	36

<i>OBSAH</i>	3
6 Eye movements in repeated multiple object tracking	38
7 Gaze position lagging behind scene content in multiple object tracking: Evidence from forward and backward presentations	55
8 Flipping the stimulus: Effects on scanpath coherence?	69
9 Scene categorization in the presence of a distractor	82
10 Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning?	94
11 Perception of direct vs. averted gaze in portrait paintings: An fMRI and eye-tracking study	109
12 Illuminating Smiles and Frowns: Visual-Affective Cueing Influences Viewer Perceptions of Page Layout Images	122
13 Tracking artistic transformations: comparing paintings and their source photographs	146
14 Podíl autora na přiložených studiích	164
15 Literatura	166

Předmluva

Předložená práce představuje mé původní studie zabývající se experimentálním výzkumem pozornosti a očních pohybů. Cílem bylo představit jeden užší výsek mé výzkumné práce. Při výběru bylo nutné řadu studií vynechat — ve výběru tak například chybí studie o paměti na fotografie, vývoji diagnostických metod či trénování paměti.

Studie jsem vybíral podle několika kritérií. Zaprvé tematicky — jedná se o studie zabývající se pozorností. Z druhé metodologicky — jedná se o experimentální výzkum a všechny předložené studie také spojuje metoda sledování očních pohybů. Zatřetí podle typu publikace — všechny studie byly publikovány ve vědeckých časopisech s impakt faktorem (až na jednu výjimku zahraničních). První polovina studií jsou metodologické publikace, které tvoří těžiště práce. Metodologické výsledky dále diskutuji na příkladech studií z druhé části, kde představuji aplikované použití sledování očních pohybů pro výzkum dalších otázek.

Celkově se jedná o osm studií z posledních deseti let mé práce. V textu současně zmíním i výsledky dvou dalších našich studií, které se věnovaly technickým aspektům metody a byly publikovány v recenzovaných zahraničních sbornících. Tento text chápu jako příležitost téma a jednotlivé výzkumy souhrnně představit v českém jazyce.

Úvodní část této práce představuje výzkumný záměr (Kap. 1) a základní teoretická a metodologická východiska (Kap. 2). Je zde popsána problematika dělení pozornosti na více cílů a metoda sledování více objektů (Multiple Object Tracking). V Kap. 3 představuji jednotlivé metodologické studie, jejich výsledky a přínos. Celkové výsledky jsou v bodech shrnuty v Kap. 4. Na závěr demonstruji metodologické důsledky na dalších příkladech (Kap.5). Jednotlivé články jsou uvedeny jako samostatné kapitoly (Kap. 6-13).

V Praze, říjen 2021

Jiří Lukavský

Kapitola 1

Úvod

Často potřebujeme sledovat více věcí v našem okolí. Řízení auta vyžaduje, abychom si uvědomovali, kde jsou nejbližší účastníci silničního provozu. V týmových sportech si potřebujeme pamatovat, kde jsou naši spoluhráči a kde soupeři. Na koupališti potřebujeme sledovat, kde jsou naše děti.

Naše oči jsou velmi dobře vyvinuté k vnímání středu zorného pole. Kvalita informací na periférii klesá a pokud nás nějaký předmět zajímá, musíme na něj nasměřovat pohled. Současné sledování více objektů tedy vyžaduje rozhodování, kdy se blíží nějaká riskantní situace a pohled na ni potřebujeme přesunout a kdy se můžeme spolehnout na méně spolehlivé informace z periferie zorného pole.

Jak tedy vlastně dokážeme zaměřit pozornost na více věcí současně? Pochopení této otázky může v budoucnu umožnit praktické aplikace. S pokračujícím rozvojem technologií bude sledování očních pohybů mnohem přístupnější a možná bude součástí řady činností. Pokud budeme sledovat oční pohyby řidiče auta nebo leteckého operátora, možná dokážeme odhadnout, kdy je nějaký objekt přehlížený a jemně na něj řidiče upozornit. Naopak v případech, kdy z chování a ze směru pohledu předpokládáme, že řidič o objektu ví, nebude systém řidiče rušit.

1.1 Odborný zájem

Mým hlavním odborným zájmem je otázka, jak lidé pozorují objekty ve svém okolí a jak se to projevuje v zaměření jejich očních pohybů. Zajímá mne, jak se při tom rozhodují i jak robustní nebo naopak náhodné tyto projevy jsou a co je ovlivňuje.

Tato práce má metodologický charakter. Cílem není popsat chování v konkrétních situacích: jak a kam se lidé dívají při sledování filmu, řízení nebo hlídání dětí. To jsou poměrně komplexní situace, kde by bylo snadné nechat se zahltit detaily. Většina předložených metodologických studií se zabývá modelovou situací zvanou Multiple Object Tracking (MOT, viz Kap. 2.2). Výhodou modelu je, že umožňuje odhlédnout od vizuálních detailů (vzhled a rozdíly mezi objekty, jejich odlišení od pozadí) a je možné tak lépe studovat samotné zaměření pozornosti.

Přestože jsou v této práci porovnávány různé dosud publikované modely očních pohybů nebo je navrhován vlastní model (anticrowding, Lukavský, 2013), cílem není formulovat jeden finální model, ale vytvořit nástroje, které umožní formulaci a testování budoucích modelů. Zabývám se tím, jak v dynamických úlohách efektivně porovnávat oční pohyby mezi sebou a také s předpovědí modelu. Zaměřuji se na celková omezení modelu – jak nejlépe můžeme oční pohyby předpovědět. A testuji vliv dalších faktorů, jako je čas nebo přítomnost nerelevantních objektů z okolí, abychom věděli, zda či jakým způsobem by měly být do modelů zahrnovány.

V metodologických závěrech zůstávám většinou na úrovni modelu (kam se lidé dívají při MOT), ale nástroje a vyzozorovaná omezení jsou obecnější a je možné je použít i v dalších situacích jako by bylo například sledování filmu. Čtvrtá zařazená studie (Lukavský, 2019) ilustruje studium vlivu rušivých podnětů při sledování fotografií.

Následující Kapitola 5 ukazuje, jak se tyto metodologické výsledky projeví na reálném výzkumu očních pohybů. Předkládané aplikované studie jsou klasickými příklady využití metody, primárně nestaví na zjištěních z metodologické části, ale dovolí nám ukázat na omezení, se kterými se výzkum potýká.

1.2 Zaměření jednotlivých metodologických studií

Záznamem očních pohybů při sledování dynamické scény získáváme velmi komplexní datové záznamy o poloze a čase. Nejen v tom smyslu, že je jich velké množství, ale také proto, že jejich vzájemnou podobnost či rozdílnost můžeme hledat na několika rovinách. Studie prezentované v této práci se postupně zaměřují na hlavní aspekty takového porovnávání. Studoval jsem možnosti **celkového porovnávání časoprostorových dat**, porovnávání **polohy pohledu** s předpovědí, **časovým vztahem** mezi pohledem a prostředím a vlivem **dalších objektů**, které nejsou pro daný úkol relevantní.

Experimenty v představených metodologických studiích se tak zaměřují na následující otázky:

- a) Dá se z očních pohybů během opakované prezentace úlohy odhadnout robustnost očních pohybů, tj. jaké množství pohledů je záměrné a účelné, a jaké náhodné (Lukavský, 2013)? Jakým způsobem bychom vlastně měli porovnávat oční pohyby, abychom z nich vytěžili více informací?
- b) Jakým způsobem bychom měli statisticky testovat podobnost očních pohybů v různých podmínkách? Jak odhalit, zda míra odchylky přesahuje obvyklou míru variability (Děchtěrenko et al., 2017)?
- c) Do jaké míry předpovídáme pohyb na obrazovce a přizpůsobujeme tomu své pohyby očí (Lukavský & Děchtěrenko, 2016)? Pokud budeme formulovat modely o směru pohledu v proměnlivé situaci, měli bychom vycházet z aktuálního stavu, nebo raději z minulého či budoucího?
- d) Dokážeme ignorovat podněty, kterých si nemáme všimnout (Lukavský, 2019)?

Celkově si tak předkládaná habilitační práce klade otázku:

Jak směřujeme svůj pohled v situacích, kdy potřebujeme sledovat více cílů současně? Jaké faktory směr pohledu ovlivňují? Můžeme si z těchto výzkumů dynamických situací přinést něco do obecného výzkumu očních pohybů?

Kapitola 2

Teoretický kontext

2.1 Pozornost

Lidská mysl nezpracovává a ani nedokáže zpracovat všechny signály, které registruje. Oba tyto aspekty (výběrovost a omezené zdroje) tvoří základ, který se psychologické teorie pozornosti snaží popsat a vysvětlit. Řada výzkumů se odvolává na intuitivní představu o tom, co do pozornosti patří, podobně jako William James (1890): „Každý ví, co je to pozornost. Mysl se zmocňuje jednoho z několika simultánních běhů myšlenek či možných objektů v čisté a jasné podobě. Jejím základem je zaměření, koncentrace vědomí. To znamená nutnost odstoupení od některých věcí, abychom se mohli účinně zabývat jinými.“¹

Výzkum pozornosti za posledních sto let přinesl řadu poznatků a pokračuje i nadále. Shrnout tyto poznatky by bylo nad rámec této práce, ale rád bych se zastavil u kritických ohlasů, které se v posledním desetiletí objevují a upozorňují na problémy tohoto výzkumu.²

Anderson (2011) upozorňuje, že už Jamesova definice v sobě ukrývá potenciální problém, protože není jasné, zda si pozornost představujeme všichni stejně. Přestože základní představa o funkci pozornosti je stejná, v učebnicích najdeme řadu dalších specifických funkcí (Hommel et al., 2019). Pozornost tak považujeme za schopnost vybírat události pro další zpracování (focused attention); schopnost ignorovat zavádějící nebo nepodstatné signály (selective attention); mechanismus integrující informace různých modalit patřící k jednomu objektu (feature integration); upřednostňování signálů z konkrétního místa v prostoru (spatial attention); systematické vyhledávání (visual search); schopnost provádět několik činností najednou (divided attention); ovládání očních pohybů (selective attention for action); výběr z nabídky možných cílů (goal-centered attention); upřednostňování jednoho objektu, vzpomínky či představy před ostatními (object-centered attention); soustředění na příchod očekávané události (sustained attention).

Současná kritika výzkumu pozornosti (Anderson, 2011; Hommel et al., 2019; Krauzlis et al., 2014) se dá shrnout do dvou základních bodů. Prvním je nejasné vymezení pozornosti. Jak již bylo naznačeno, do pozornosti

¹Překlad Dalibor Špok (Eysenck & Keane, 2008) a vlastní překlad.

²Kritika míří k samotné podstatě pojmu pozornosti, jak je patrné i z názvů studií: „There is no Such Thing as Attention“ (Anderson, 2011) nebo „No one knows what attention is“ (Hommel et al., 2019)

nyní zahrnujeme celou řadu funkcí. Tyto funkce se často překrývají a je obtížné je mezi sebou přesně rozlišit. Výzkum se dosud často soustředil na definování dichotomií (bezděčná a záměrná pozornost; dvě stádia vyhledávání), ale tyto dichotomie nejsou příliš ostré a často popisují spíše kvantitativní než kvalitativní rozdíl ve fungování pozornosti. Kromě tohoto problematického vymezování v rámci pojmu narážíme na obtížné oddělování pozornosti od dalších funkcí, jako je rozhodování či motivace.

Druhým problémem je způsob výkladu, který v souvislosti s pozorností používáme. Často není jasné, zda pozornost chceme vysvětlit z pozorovaných jevů nebo naopak zda chování vysvětlujeme pozorností (např. reprezentuje pozornost omezenou kapacitu naší kognice nebo způsob, jak se s omezenou kapacitou vyrovnat?). Dlouhodobě převládá kauzální výklad, kdy pozornosti přisuzujeme existenci (reifikace), mluvíme o „úzkém hrdle“ či „reflektoru“ a používáme tyto metafory k výkladu příčin chování.

Kiritici proto navrhuji obrácení kauzálního výkladu a o pozornosti mluvit spíše jako o důsledku chování (Anderson, 2011; Krauzlis et al., 2014). Roli pozornosti je možné například přeformulovat na bayesovský proces rozhodování, kdy se člověk rozhoduje v nejisté situaci na základě předchozích zkušeností, odhadovaných pravděpodobností a očekávaných odměn (Anderson, 2011). Hommel a kol. (2019) navrhuji opustit současnou rovinu výkladu, nemluvit o pozornosti jako psychickém mechanismu nebo funkci mozku, ale spíše mluvit o konkrétních procesech a chování, které se selektivitou souvisí.

Ve svém výzkumu se převážně zabývám zrakovou pozorností — chováním, kdy lidé po určitou dobu sledují několik pohybujících se objektů. V souvislosti se specifickými funkcemi pozornosti zmíněnými výše bychom mohli mluvit o *visual sustained distributed attention*. Za velkou výhodu tohoto konkrétního přístupu považuji modelovou situaci (Multiple Object Tracking, Kap. 2.2). O úloze budu podrobně mluvit dále, zde jen zmíním, jak úloha může vypadat — např. cílem je sledovat 4 pohybující se objekty mezi 4 dalšími stejnými objekty po dobu 10 vteřin a pak je správně identifikovat. V porovnání s jinými přístupy (např. sledovat fixační křížek uprostřed obrazovky a rozpoznat nejasný objekt promítнутý na okraji zorného pole)³ je úloha komplexnější, ale současně názorná a úspěch či neúspěch v ní je pro účastníky očividný. Chování člověka během úlohy můžeme analyzovat jako celek. V úloze nepochybně hrají roli automatické procesy i strategická rozhodnutí, ale protože se zaměřuji na pozorovatelný projev chování (oční pohyby), není potřeba příliš jednotlivé mechanismy vydělovat.

Předložené studie jsou psané s ohledem na obvyklý současný diskurz předpokládající pozornost jako samostatnou psychickou funkci. Podobně v následující kapitole o rozdělené pozornosti budu navazovat na obvyklé metafory pozornosti („reflektor“). Samotná výzkumná otázka vlivu vzdálenosti, času, šumu či irelevantních objektů na chování však není na konkrétním teoretickém rámci příliš závislá.

2.2 Sledování více objektů

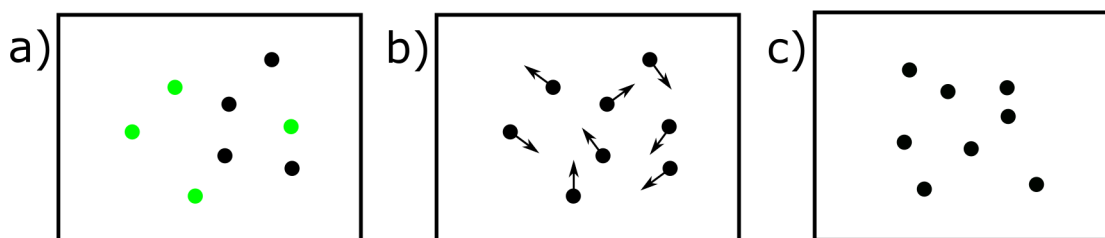
Teorie vizuální pozornosti i obvyklé experimentální přístupy vycházejí z představy pozornosti, která se zaměřuje v konkrétním okamžiku na jedno místo nebo objekt. Problematika vizuální distribuované pozornosti (zaměření na více míst nebo objektů) nebyla dlouho zkoumána a její rozvoj souvisí s metodou sledování více objektů

³Zjednodušený popis tzv. attentional cueing (Posner, 1980).

(Multiple Object Tracking, MOT). Metodu navrhli Pylyshyn a Storm (1988) a od té doby se v různých obměnách používá ve výzkumu dodnes. Ukázkou úlohy najdete na Obr. 2.1. Na začátku každé úlohy účastník vidí několik stejných objektů (např. 8). Některé z objektů jsou na krátkou chvíli označeny (změní barvu, zablikají). Těmto objektům se říká cíle (*targets*) a úkolem účastníka je tyto cíle sledovat. Cíl může být jeden, ale obvykle jich bývá víc (např. čtyři). Zbývající objekty se nazývají distraktory (*distractors*).

Když označení zmizí, objekty vypadají opět identicky a začnou se po obrazovce pohybovat. Po krátké chvíli, trvající obvykle 7-12 vteřin, se zastaví a testujeme, zda si účastník udržel přehled o cílových objektech. To vypadá tak, že buď myší označí všechny cílové objekty, a nebo je označen jeden z objektů a účastník má rozhodnout, zda se jedná o cílový objekt nebo ne.

Odpověď úloha končí a následuje další podobná úloha. Jednotlivé úlohy z pohledu účastníka vzájemně nesouvisí a nemusí si z nich po jejich skončení nic dalšího pamatovat. Experiment obvykle obsahuje desítky podobných úloh (včetně kratších přestávek).



Obrázek 2.1: Struktura MOT úlohy. (a) Nejprve je zvýrazněno, které objekty má účastník sledovat. (b) Poté se objekty stanou nerozlišitelné a začnou se pohybovat. (c) Nakonec se zastaví a účastník má vybrat původně označené objekty.

Úloha MOT byla původně vyvíjena pro výzkum pojmu objekt (Pylyshyn, 2007). Cílem bylo popsat, jak lidé vydělují ze své zkušenosti související vizuální signály a považují je za součást jednoho objektu. Předmětem výzkumu tak byly například otázky, zda lidé dokáží sledovat koncové body úseček nebo dvojice propojených objektů (Scholl et al., 2001; Scholl, 2009).

Velký zájem o MOT byl podnícen teoretickou otázkou, jak pozornost obvykle popisovaná jako „úzké hrdlo“ či „reflektor“ a zaměřená na malý výsek okolního prostoru, může sloužit v takto komplexní úloze vyžadující přehled o různých částech scény. Teoriím MOT se věnuji v další kapitole.

V poslední době přibývá studií, které hledají interindividuální rozdíly ve výkonu v této úloze. Většinou se jedná o otázky, zda se některé skupiny odlišují. Zkoumá se tak výkon sportovců (Howard et al., 2018), operátorů v letectví (Allen et al., 2004), seniorů či hráčů počítačových her (Sekuler et al., 2008). Úloha je také využívána pro měření efektivity intervencí. Přestože je tento zájem pochopitelný, používaný obvyklý design MOT je vhodnější spíše pro experimentální výzkum a pro diagnostické účely není příliš efektivní. Je zde tedy stále prostor pro rozvoj metody a její adaptaci pro diagnostiku a klinické aplikace.

2.2.1 Teorie MOT

Otázkou, jak zaměřená pozornost může sloužit sledování více různých objektů, se zabývala již první studie (Pylyshyn & Storm, 1988), kde autoři definují tzv. **sériový model** MOT⁴ a na základě dosud známých parametrů pozornosti jej vyvrací. Podle sériového modelu si v úvodní fázi vytvoříme seznam cílů a jejich pozic na obrazovce. V průběhu sledování přesouváme pozornost postupně podle pozic v seznamu. Doba přesunu pozornosti je závislá na vzdálenosti mezi pozicemi. Protože se objekty pohybují, zapamatované pozice již nejsou po přesunu aktuální, a proto vyhledáme nejbližší objekt. Aktualizujeme si pozici v našem seznamu a přesouváme se na další objekt. Takto cyklicky navštěvujeme všechny cíle, dokud se objekty nezastaví. Nejrizikovější fází popsaného procesu je dohledávání nejbližšího objektu po přesunu pozornosti, kdy předpokládáme, že to je objekt, který se původně nacházel na uložené pozici. Toto dohledávání je těžší s rostoucím počtem sledovaných cílů, protože trvá déle, než objekt znovu navštívíme, a tak se může posunout dál. Rovněž další faktory jako rychlost nebo počet distraktorů dohledávání komplikují. Pylyshyn a Storm (1988) ukázali, že sériový model předpovídá mnohem nižší úspěšnost, než ve skutečnosti u lidí pozorujeme

Jako alternativu k nevyhovujícímu sériovému modelu navrhli Pylyshyn a Storm vlastní teorii vizuálních indexů pojmenovanou **FINST**⁵. Jedná se o obecnější teorii (Pylyshyn, 1989, 2001), která se zabývá vztahem mezi reprezentací objektů v naší mysli a vstupními vizuálními údaji (na úrovni sítnice). Pylyshyn předpokládá, že existuje omezený počet ukazatelů (*visual index*), které odkazují na určitá místa zrakového pole a umožňují nám chápat údaje z toho místa jako objekt. Dále se předpokládá, že existují mechanismy, které automaticky aktualizují polohu těchto indexů, pokud se objekt posune nebo pohneme očima. Teorie FINST je obecnější teorie, aspirující na popis zrakového vnímání objektů, která se uplatňuje i pro popis dalších jevů, jako je zrakové vyhledávání nebo subitizing⁶ (Pylyshyn, 1994). V rámci MOT tato teorie tvrdí, že každý z cílů má přiřazený svůj ukazatel a automatické mechanismy mozku se postarají o jeho aktualizaci během pohybu. V případě potřeby indexy umožňují rychlý přesun pozornosti na odkazované místo nezávisle na vzdálenosti.

Yantis (1992) navrhnul, že lidé cíle nesledují jako individuální objekty, ale seskupí si je v mysli (**grouping model**) do rovinného útvaru. Úloha se pro ně potom změní ve sledování geometrických deformací tohoto trojúhelníku či čtyřúhelníku. Účastníci experimentů často popisují, že na vědomé úrovni takto postupovali. Ve prospěch tohoto modelu také hovoří další nepřímé důkazy. V první řadě, je snazší sledovat objekty, které udržují během pohybu stálý celkový tvar (Yantis, 1992). Za druhé, lidé se při MOT často dívají do blízkosti středu celkového tvaru (centroid). Za třetí, EEG studie ukazují specifické reakce na drobné podněty zobrazované uvnitř a na hranách celkového obrazce (Merkel et al., 2017).

Dosavadní modely předpokládaly pozornost, která se zaměřuje na jediné místo. V závislosti na daném modelu se pozornost přesouvala k jednotlivým cílům nebo sledovala virtuální útvar, který objekty tvoří. Model víceohniskové pozornosti (**multifocal attention model**, Cavanagh & Alvarez, 2005) předpokládá, že disponujeme několika ohnisky pozornosti a při MOT je každý z cílů sledován jedním takovým ohniskem. Návrh sice

⁴Také nazývaný jako *switching model* (např. Cavanagh & Alvarez, 2005)

⁵Název *FINST* je odvozen z anglického "FINgers of INSTantiation"

⁶Subitizing označuje lidskou schopnost rychle určit počet zobrazených objektů. Malé počty do 4 objektů zjistíme okamžitě (v konstantním čase), při větších počtech se potřebný čas lineárně prodlužuje s počtem počítaných objektů.

připomíná indexy FINST modelu, podstatný rozdíl je ale v tom, že zde jsou všechny cíle neustále sledovány pozorností, zatímco podle FINST modelu může být pozornost jen rychle alokována k jednomu libovolnému cíli. Hlavním dokladem pro tento model jsou pozorované kapacitní limity v rámci zorného pole. Naše kapacita se zmenší na polovinu, pokud jsme nuceni sledovat cíle v rámci jedné poloviny zorného pole oproti situaci, kdy se cíle nachází v obou polovinách (Alvarez & Cavanagh, 2005; Carlson et al., 2007). Obrazy z jednotlivých polovin zorného pole jsou zpracovávány zvlášť, každý v jiné mozkové hemisféře. Jinými slovy, místo schopnosti sledovat 4 pohyblivé cíle jsme spíše schopni sledovat 2 cíle v levé polovině a 2 cíle v pravé polovině zorného pole.

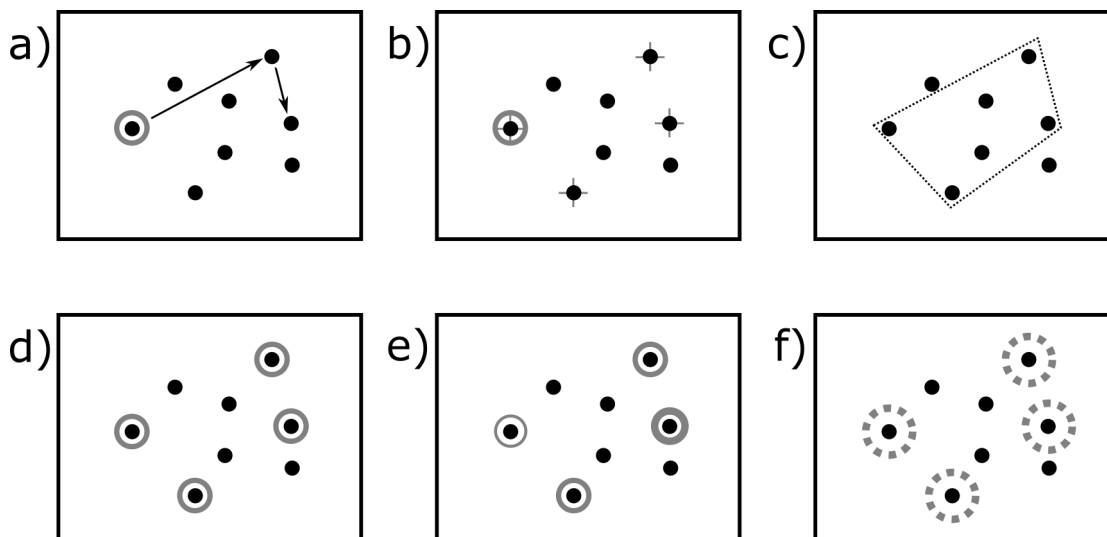
Jedním z aktuálních témat psychologie je problém, zda je naše kognitivní kapacita rozdělována po diskretních celcích (slot model) nebo zda může být distribuována spojitě dle nároků úlohy (resource model). Tato debata se objevuje například v teoriích pozornosti nebo krátkodobé paměti (Suchow et al., 2014). Volně řečeno, první přístup říká, že máme určitý omezený počet jednotek, které odpovídají prvkům, které si můžeme pamatovat. Pokud si potřebujeme zapamatovat komplexnější objekt, můžeme na jeho zapamatování použít více jednotek. Druhý přístup předpokládá, že existuje pouze celková dostupná kapacita, kterou můžeme dělit podle potřeby — tak si můžeme pamatovat větší počet jednoduchých prvků nebo naopak. Tato debata není rozhodnutá a obsahuje mnoho otevřených otázek (Suchow et al., 2014).

Pozorování, že lidé jsou schopni sledovat přibližně 4 pohyblivé cíle, by odpovídalo diskretnímu modelu. Myšlenku spojitě distribuovatelné pozornosti lze ale aplikovat i na MOT. Alvarez a Franconeri (2007) manipulovali rychlostí pohybu objektů a ukázali, že v případě velmi malých rychlostí lidé dokáží sledovat až 8 cílů. Ve svém **FLEX modelu**⁷ (Alvarez & Franconeri, 2007) navrhují, že lidé disponují vizuálními ukazateli pro paralelní sledování. Jejich počet není omezen, nicméně s počtem indexů klesá jejich účinnost, protože čerpají ze společného zdroje. Myšlenku proměnlivé kapacity také podporují pozorování, že si přesněji pamatujeme cíle, u nichž bylo větší nebezpečí záměny s jiným objektem (Iordanescu et al., 2009). Problémem FLEX modelu je neurčitost jeho předpovědi — pozorování je možné dodatečně vysvětlit úpravami předpokladů modelu. Model také nezohledňuje pozorování týkající se nezávislé kapacity pro sledování v jednotlivých částech zorného pole, které jsme zmiňovali u modelu víceohniskové pozornosti.

Výkon v MOT ovlivňuje řada parametrů, jako je počet cílů či distraktorů, jejich rychlost či doba trvání úlohy. Franconeri a kol. (2010) upozorňují, že pro vznik chyb je kritický okamžik, kdy jsou objekty blízko sebe a ostatní zmíněné faktory s kolizemi souvisí. Jinými slovy, větší počet objektů vede k většímu počtu vzájemných srážek a rovněž delší trajektorie objektů (ať již díky vyšší rychlosti či delšímu času) s sebou přináší více srážek. V souladu s tím navrhli **interferenční teorii** MOT (spatial interference theory), která popisuje distribuci pozornosti v okolí sledovaných objektů. Teorie v podstatě staví na předchozí představě víceohniskové pozornosti, kterou doplňuje o zjištění, že dochází k potlačování signálů z okolí pozorovaných objektů (např. Hopf et al., 2006). Signály ze sledovaného objektu a jeho nejbližšího okolí jsou zesíleny (proto dochází k chybám při srážce cíle s distraktorem). Za tímto okolím je oblast, kde jsou signály potlačovány. To vysvětluje, proč i blízkost dvou cílů vede ke zhoršení výkonů (Shim et al., 2008).

Z výše uvedeného popisu je zřejmé, že v tuto chvíli existuje několik teorií pro popis pozornosti v průběhu MOT.

⁷Název *FLEX* je odvozen z anglického “a Flexibly allocated indEX”



Obrázek 2.2: Teorie MOT. (a) Podle **sériového modelu** se pozornost postupně zaměřuje na všechny sledované objekty. (b) **FINST model** předpokládá existenci odkazů, které sledují objekt, a pozornost se na ně může snadno přesunout. (c) Pokud si lidé **seskupí** objekty do složitějšího obrazce, mohou místo sledování několika objektů pozornost zaměřit na geometrické proměny jednoho obrazce. (d) **Víceohniskový model** předpokládá, že dokážeme pozornost zaměřit na několik objektů současně. (e) **FLEX model** považuje pozornost za spojitý zdroj, který je možné dle potřeby distribuovat mezi sledované objekty. (f) **Interferenční teorie** tvrdí, že každý sledovaný objekt je současně obklopen zónou, která potlačuje signály z jeho okolí.

Jednotlivé teorie sice disponují jedinečnými pozorováními, která je podporují, ale současná úroveň metod neumožňuje přesně rozlišovat mezi predikcemi jednotlivých teorií. Jedním z problémů je, že informace o alokaci pozornosti jsme schopni získávat jen nepřímo, např. z reakcí na rychle zobrazené objekty v zorném poli. Výraznější nasazení takové úlohy ale může narušovat průběh původní úlohy. Pokud se během MOT budeme ptát na větší množství irelevantních podnětů, lidé na ně začnou dávat pozor a původní úloha sledování MOT se z jejich perspektivy změní na sledování MOT při současné detekci objektů v zorném poli.

2.3 Oční pohyby při MOT

Další možností, jak studovat alokaci pozornosti, je měřit oční pohyby. Předem je potřeba zdůraznit, že předmět, na který se díváme, nemusí být předmět, na který dáváme pozor. V rušné místnosti můžeme s někým konverzovat a dívat se mu do očí a současně dávat pozor, zda se neotevřou dveře nebo zda nás nemíjí někdo další. Rovněž tradiční experimenty zaměřené na studium pozornosti ukazují, že zaměření pohledu a pozornosti jsou vzájemně disociovatelné. Například se můžeme dívat doprostřed obrazovky, kde se zobrazuje šipka a podle směru šipky dávat pozor na objekty v levé/pravé části periferního zorného pole (Posner, 1980).

Obecně při měření očních pohybů pozorujeme krátká období, kdy je oko v relativním klidu a zaměřené na konkrétní místo. Těmto obdobím klidu se říká *fixace*⁸ a obvykle trvají 200-300 ms, i když jejich délka může

⁸Termín *fixace* se používá i pro místo pohledu, kam je oko v tomto období zaměřeno.

být závislá na konkrétní úloze a její obtížnosti. Oko se mezi jednotlivými fixacemi přesouvá rychlými skoky, kterým se říká *sakády* (saccade). Sakády jsou kratší než fixace, trvají asi 70 ms v závislosti na délce přesunu. Během sakadických pohybů je naše schopnost vnímat svět okolo výrazně utlumena. Plánování očních pohybů s sebou tedy přináší nutnost volby. Přesun pohledu na nové místo nám umožní získávat podrobnější informace, ale současně riskujeme, že během přesunu nebude schopni reagovat na změny v okolí. Toto dělení na fixace a sakády nacházíme ve většině situací. Kromě těchto období klidu a prudkých skoků pozorujeme i chvíle, kdy se oko pohybuje plynule. *Plynulého sledování* (smooth pursuit) jsme schopni v případě, že náš pohled je zaměřen na pohybující se objekt v našem zorném poli.

V raných MOT experimentech (např. Pylyshyn & Storm, 1988) byli účastníci žádáni, aby nehýbali očima, upřeně se dívali na střed obrazovky a pohybující se objekty sledovali pouze pozorností. Základem prvních systematických studií očních pohybů během MOT (Fehd & Seiffert, 2008, 2010; Zelinsky & Neider, 2008) byla analýza směru pohledu během jednotlivých snímků animace. Podle místa pohledu a polohy objektů na obrazovce pak byly tyto úseky klasifikovány. Tímto způsobem uvedené studie odhalily dvě základní strategie: sledování cílů a sledování centroidu⁹. Volba konkrétní strategie záleží mj. na počtu sledovaných objektů. V triviálním případě, pokud máme sledovat jeden objekt, tak jej sledujeme i pohledem skoro neustále (96%), zatímco při sledování 3 objektů se díváme na jednotlivé cíle již méně často (26%) a nejčastěji se díváme do blízkosti centroidu (66%, Fehd & Seiffert, 2008). Podle dalších studií (Zelinsky & Neider, 2008) se lidé při sledování 2 cílů dívají častěji do blízkosti centroidu než na jednotlivé cíle. Při sledování 3 nebo 4 objektů jsou obě strategie podobně časté. Zároveň mírně roste tendence dívat se na jednotlivé cíle u vyššího počtu sledovaných prvků. To může být dáno potřebou podívat se na vzdálený cíl, který se blíží do nebezpečné situace (kolize). Tomuto očnímu pohybu se v literatuře říká „rescue saccade“ (Zelinsky & Todor, 2010).

Procentuální hodnoty napříč jednotlivými studii nejsou přímo srovnatelné, protože záleží na vizuální podobě zobrazených podnětů a rovněž na přesných klasifikačních kritériích. První studie kategorizovaly pohled podle toho, které místo bylo nejbližší (cíl, distraktor, centroid). Nevýhodou tohoto přístupu je, že úplně každý pohled je klasifikovaný, bez ohledu na to, jak dobře odpovídá cílovému místu. Pozdější studie se zdokonalenou metodologií klasifikace (Huff et al., 2010) již používala tzv. oblasti zájmu (area of interest, AOI). Zde musel být pohled i dostatečně blízko od cílového místa (cíl, centroid), aby byl klasifikován. Studie potvrdila důležitost sledování centroidu, ale zároveň víc než 50 % pohledů nebylo klasifikováno ani jako cíl, distraktor nebo centroid.

2.4 Modelování očních pohybů

V rámci výzkumu pozornosti se vědci snaží formulovat modely očních pohybů. Nejúspěšnější jsou v tomto ohledu tzv. *modely salience*, které od původního modelu (Itti et al., 1998) prošly bohatým rozvojem. Tyto modely na základě analýzy fotografie předvídají, na které části se budou lidé dívat. Modely mají řadu využití od ergonomie (testování výraznosti dopravních značek nebo naopak efektivity kamufláže) po automatické generování náhledů fotografií tak, aby obsahovaly pro lidi nejdůležitější informace. Pro své předpovědi modely

⁹Centroid může být definován různými způsoby, nejčastěji se jedná o fyzikální těžiště sledovaných bodů (tj. izolovaných bodů, ne obrazce, který tvoří). V textu budu používat původní pojem centroid.

analyzují nízkourovňové parametry fotografie (barva, jas, orientace linií) a jejich distribuci v rámci celku. Nové modely rovněž zohledňují přítomnost lidských postav či textu (Bylinskii et al., 2016), pracují s odhadovanou mírou informativnosti (Henderson & Hayes, 2017) nebo staví na nových metodách strojového učení (Kummerer et al., 2017).

Pro naše účely bych chtěl shrnout dva aspekty těchto úspěšných modelů, které souvisí s reliabilitou. Zaprvé, oční pohyby při prohlížení fotografie můžeme od každého člověka nahrát pouze jednou a nemáme tak přirozené srovnání, jak stabilní je chování, které chceme předpovídat. Při měření očních pohybů má velký význam úkol, který člověk dostane (DeAngelus & Pelz, 2009; Yarbus, 1967), a proto jej výzkumníci volí velmi opatrně. V případě výzkumů salience většinou účastníky požádají, aby si fotografie prostě prohlíželi bez dalšího cíle (free viewing), nebo aby se pokusili fotografií co nejlépe zapamatovat (recognition), případně na fotografii něco našli (visual search). Pokud bychom účastníkům ukázali stejnou fotografii podruhé, ani při jedné z těchto instrukcí nebudou lidé oční pohyby opakovat stejným způsobem. Možná se budou vracet se k tomu, co je zaujalo, ale také budou zkoumat části, které při prvním prohlížení nestihli, případně najdou hledaný předmět mnohem rychleji. I když čteme stejný text podruhé, projeví se to na očních pohybech (Raney & Rayner, 1995). Modely nám tedy umožňují určitou míru predikce, ale díky absenci dat o intraindividuální variabilitě chybí srovnání, jakou míru úspěšnosti resp. šumu můžeme čekat.

Zadruhé, výsledkem modelů je jedna distribuce pravděpodobnosti pohledu. Modely pracují tak, že agregují data od mnoha lidí, což jim umožňuje odhadnout, které pohledy lze pravděpodobně považovat za šum a které informace zaujaly více lidí. Ale právě díky této agregaci zobecňují a kladou důraz na univerzální vlastnosti obrazů i lidí. Existují přitom doklady o tom, že se lidé liší v očních pohybech a tyto odlišnosti se dají alespoň částečně vysvětlit mezilidskými rozdíly, jako jsou kognitivní a osobnostní vlastnosti, genetika nebo předchozí zkušenost. Pro ilustraci uveďme pár příkladů. Lidé s lepší pracovní pamětí dělají delší sakadické pohyby při čtení a delší fixace při prohlížení fotografií (Luke et al., 2018). Zvědaví lidé při prohlížení prozkoumají větší plochu obrazu (Risko et al., 2012). Distribuce pohledů jsou podobnější u monozygotních dvojčat (Kennedy et al., 2017). Existuje řada výzkumů dokumentujících rozdíly v očních pohybech s rostoucí mírou doménové zkušenosti, např. u lékařů nebo šachistů (Reingold & Sheridan, 2011). V našem vlastním výzkumu (Adámek et al., 2019) jsme porovnávali, zda se oční chování laiků a odborníků na výtvarné umění¹⁰ liší při prohlížení výtvarných děl (malby, fotografie) a zjistili jsme, že odborníci prozkoumávali větší plochu díla.

Pokud se lidé při nějaké úloze dívají na určitá místa, může to mít různé důvody. Může se jednat o různě komplexní motivy od (1) salientních charakteristik obrazu, přes (2) význam objektů a (3) plnění úkolů a plánů po (4) očekávanou odměnu (Schütz et al., 2011). Pomiňme pro naše účely vliv odměny. Pokud analyzujeme oční pohyby při prohlížení fotografie či videa, stávající modely zdůrazňují zejména první dvě roviny, i když zadaný úkol může hrát také roli. U MOT sledujeme velmi jednoduché geometrické tvary, proto význam objektů není relevantní, a hlavní motivy budou z třetí a také první roviny — tedy náš plán a geometrické vlastnosti situace.

V první řadě, při MOT má účastník jasný úkol sledovat objekty. Úkol trvá po celou dobu pohybu a není jej možné splnit předčasně. Člověk může na základě průběhu svého výkonu a související míry jistoty/nejistoty plánovat své oční pohyby. Například, pokud bude hrozit kolize objektů na okraji zorného pole, může člověk

¹⁰Výtvarníci, historici umění a pokročilí studenti těchto oborů.

dočasně toto místo zkontrolovat pohledem, aby zmenšil nebezpečí záměny identit objektů (rescue saccade).

Konfigurace objektů na obrazovce a její změny mohou zvýhodňovat některá místa pohledu. Fixace centroidu může být jedním takovým příkladem. Ale konfigurace může ovlivňovat pohled i v dalších situacích. Pokud objekty vytvoří malý hlouček, může to vést k pohledu do středu takové skupinky, podobně jako lidé mají tendenci dívat se do středu objektů (Nuthmann & Henderson, 2010). Snaha minimalizovat situace, kdy dochází na periferii zorného pole ke kolizím (crowding), může být společným rysem obou motivačních rovin, záleží na tom, jakou míru automatismu u pozorovatele očekáváme.

Jedním z důsledků této úvahy je, že v prezentovaných studiích analyzujeme pouze oční pohyby ze správně zodpovězených úloh. U takových úloh můžeme předpokládat, že účastník plnil úlohu celou dobu. V situaci, že jeden sledovaný bod „ztratí“, může volit různé strategie, jako sledovat zbylé objekty nebo hledat ztracený objekt. To by přinášelo vyšší šum zejména v pozdějších fázích úlohy, kdy roste riziko, že v průběhu již ke ztrátě došlo. Určitá míra šumu bude v měření zachována i tak, protože správná odpověď může být důsledkem šťastného hádání. Obecně se snažíme prezentovat náročné úlohy, aby vyžadovaly velkou míru soustředění, ale převážně splnitelné (úspěšnost správného určení všech 4 objektů bývá alespoň 90 %), abychom zabránili zbytečným ztrátám dat.

Kapitola 3

Komentáře k prezentovaným studiím

Z předchozího shrnutí vyplývá, že jasná teorie rozdělení vizuální pozornosti zatím chybí. Některé představy sice je již možné vyloučit, ale předpovědi pozdějších teorií jsou poměrně podobné. Jednou z možností, jak získat další empirická data, je výzkum očních pohybů při sledování více pohybujících se objektů (MOT). Dosavadní studie očních pohybů při MOT mají dvě základní omezení.

Zprvé, po technické stránce spoléhají na kategorizaci místa pohledu do některé z diskrétních možností (pohled na cíl/centroid atd.). To vede k technickým úvahám, jak tolerantní máme při kategorizaci být. Pokud pozorovatel spěchá a nepotřebuje detailní informace, nemusí své oční pohyby zaměřovat přesně. Větší tolerance s sebou přináší problémy, jak kategorizovat body, které budou spadat do okruhu několika možných kategorií. Menší tolerance vede k většímu množství nezařazených pozorování. I na současné míře tolerance ale vidíme, že velká část očních pohybů není klasifikována vůbec (přibližně 50 % u Huff et al., 2010).

Zadruhé, studie vyžadují apriorní představu o tom, která místa jsou důležitá a tedy, jaké kategorie vůbec máme definovat. Může se tak stát, že některá místa výzkumníky nenapadnou. Takovými místy by mohly být třeba centroidy oddělených podskupin objektů, tj. pokud bychom chtěli ověřit, co se děje s dvojicí objektů na okraji zorného pole, možná přesuneme pohled na střed jejich spojnice. Je ale otázkou, zda je možné předem vymyslet všechny takové kategorie. Větší množství definovaných kategorií může zvýšit riziko překryvu (jak je popsáno výše) a komplikovat interpretace nebo srovnání modelů.

Dalším, obecnějším, problémem výzkumu očních pohybů je nedostatek informací o tom, jak chování rozdělit na signál a šum. Nahráváme sice záznam očních pohybů, ale u jednotlivých míst nevíme, zda byla fixována záměrně nebo zda se jednalo o bezděčné či chybné pohyby. Vlastně ani nevíme, jaký poměr mezi signálem a šumem očekávat.

V následujícím textu stručně čtenáře provádím sérií 4 studií, které se těmto problémům věnují. V první prezentované studii (**Lukavský, 2013**) představuji metodu spojitého hodnocení podobnosti očních pohybů¹. Pomocí této metody ukazují, jak lze měřit vnitřní konsistenci očních pohybů v podobných situacích a jak porovnávat

¹V této práci často používám termín “spojité hodnocení podobnosti,” protože chci upozornit, že tento přístup vyjadřuje podobnost na spojitě škále a nevyžaduje zařazování pohledů do diskrétních oblastí. Tento termín nesouvisí s termínem “spojité měření” (conjoint measurement theory), jak je používán v psychometrii.

chování s několika modely současně. V druhé studii (**Děchtěrenko et al., 2017**) se věnujeme otázce, jak statisticky porovnávat oční pohyby za dvou různých podmínek. Jako příklad si volíme srovnání očních pohybů při sledování zrcadlově převrácených scén. Třetí studie (**Lukavský & Děchtěrenko, 2016**) se systematicky zabývá porovnáváním podobnosti očních pohybů v čase. Testujeme, do jaké míry jsou oční pohyby výsledkem pozorovatelova předvídání sledovaného vývoje. V poslední studii této kapitoly (**Lukavský, 2019**) ukazují, že i ignorované a irelevantní podněty mají vliv na naše vnímání. Studie se primárně zabývá hodnocením obsahu fotografií, ale vliv nerelevantních podnětů na oční pohyby při MOT jsme ukázali i v naší jiné práci (Děchtěrenko & Lukavský, 2016).

3.1 Opakování očních pohybů

V první prezentované studii (**Lukavský, 2013**) navrhuji metodu porovnávání podobnosti očních pohybů pomocí spojitě metriky. Přesun od diskrétní klasifikace pohledu do kategorií ke spojitému hodnocení podobnosti nejen řeší výše zmíněná omezení, ale umožňuje jemnější porovnávání chování s předpověďmi modelů.

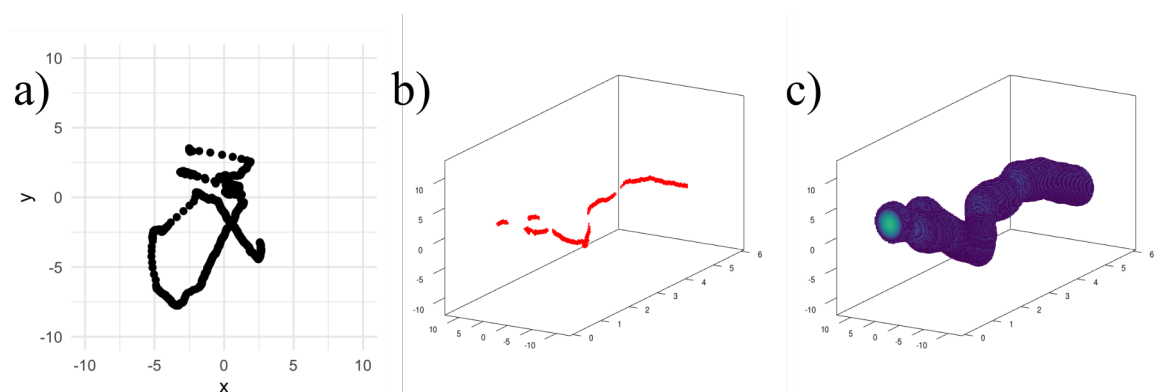
Druhým výrazným pilířem této studie je studium očních pohybů při opakovaných prezentacích pro lepší porozumění našim strategiím a rozlišení mezi signálem a šumem. Jinými slovy, pomocí opakování můžeme lépe rozlišit, které oční pohyby dělají všichni, které jsou náhodné, a třeba které jsou důsledky individuálních rozdílů. Jak již bylo zmíněno, opakované prezentace fotografií a videí jsou problematické. Na druhou stranu MOT je potenciálně dobrým kandidátem pro opakované prezentace. V pilotních experimentech jsem testoval, zda lidé poznají, že stejnou úlohu vidí opakovaně. Účastníci sice občas říkali, že měli pocit, že se úlohy opakují, ale patrně měli na mysli celkovou povahu úlohy, kterou viděli přibližně 70×. Pokud jim byly úlohy znovu ukázány a měli za úkol rozlišit, které již viděli (4×) a které jsou nové, jejich schopnost rozlišení byla na úrovni náhody. Při častém opakování velmi obtížné úlohy (sledování 5 objektů, 15×, Ogawa et al., 2009) dochází tomu, že se výkon lidí nevědomě mírně zlepšuje,² ale ani zde lidé nebyly schopni vědomě rozpoznat, které úlohy se opakovaly a které ne. V mém experimentu, kde lidé sledovali 4 objekty a úlohy se opakovaly 4× jsem nenašel doklady o tom, že by se lidé postupně zlepšovali nebo že by se podobnost očních pohybů v čase měnila (Lukavský, 2013).

3.1.1 Metodický přístup

Navržená metoda čerpá z přístupu Dorra a kol. (2010), kteří měřili variabilitu očních pohybů při sledování filmových záběrů. Postup je popsán v článku (Lukavský, 2013), ale protože se jedná o důležitou metodu i pro další publikace, krátce ji vysvětlím. Metoda spočívá v tom, že si představíme časoprostor, který reprezentuje všechna místa a okamžiky, kam se člověk během úlohy mohl dívat. Tento prostor má tři rozměry, kde první dva odpovídají souřadnicím na monitoru a třetí odpovídá času. Pro účely dalších výpočtů je potřeba tento časoprostor diskretizovat, tj. rozdělit na malé buňky, v našem případě o rozměrech $0.25^\circ \times 0.25^\circ \times 10$ ms. Stupně jsou obvyklou jednotkou měření vizuální vzdálenosti; pro představu, účastníci se dívali na monitor ze vzdálenosti

²Tomuto jevu se říká *contextual cueing* a obvykle jej pozorujeme u vyhledávání (visual search). Lidé se nevědomě naučí nalézat cílový objekt rychleji v úlohách, které se ve stejné podobě opakují (Chun & Jiang, 1998).

50 cm a tak 0.25° odpovídá přibližně 8 pixelům. Každá buňka obsahuje údaj, kolikrát se na dané místo v daném čase člověk podíval. Použitý eyetracker EyeLinkII (SR Research, Canada) nahrává polohu oka $250\times$ za vteřinu, a tak každá buňka o trvání 10 ms zahrnovala maximálně 3 záznamy. Číslo v buňce mohlo být větší, pokud jsme sloučili údaje z opakovaných prezentací nebo od různých lidí. Pokud bychom porovnávali takto vytvořené časoprostory očních pohybů, našli bychom podobnost pouze v místech, kde se člověk podíval znovu do přesně stejného místa ve stejný čas s tolerancí danou velikostí buňky. Pokud by se oční pohyby odchýlily mimo tuto toleranci, označili bychom je za odlišné a nezáleželo by, zda se odchýlily jen o pár pixelů nebo směřovaly na opačný konec obrazovky. Proto se v posledním kroku časoprostor rozostří pomocí Gaussovského filtru, umožní nám to spojitě kvantifikovat podobnost i při malých odchylkách. V našem případě jsme použili Gaussovský filtr s parametry $\sigma_x = \sigma_y = 1.2^\circ$ a $\sigma_t = 26.25$ ms. Hodnoty byly stanoveny na základě předchozí studie (Dorr et al., 2010). Hodnoty parametrů nejsou klíčové, pro účely ověření jsem výsledky přepočítával i pro poloviční a dvojnásobné hodnoty a výsledky byly analogické.



Obrázek 3.1: Zpracování dat pro srovnání pomocí Normalized Scanpath Saliency nebo Correlation Distance. (a) Rozložení očních pohybů v průběhu úlohy. (b) Převod dat do časoprostorových souřadnic. (c) Rozostření dat pomocí Gaussovského filtru.

Výše popsaný nástroj pro porovnávání se dá používat různými způsoby. V první studii (Lukavský, 2013) se jednalo o dva základní scénáře. Zaprvé, k porovnání koheze skupin očních pohybů (míra vzájemné podobnosti), např. zda se liší koheze v dané úloze u stejné osoby ve srovnání s náhodným výběrem různých osob. Vždy se jednalo o skupiny 4 záznamů. V každém kroku byl vybrán jeden záznam a porovnán s ostatními třemi, tj. byl spočítán časoprostor pro 3 záznamy a spočítáno, jak vysoké hodnoty najdeme v buňkách odpovídajícím prvnímu záznamu. Porovnání bylo zopakováno pro každý záznam (tzv. leave-one-out metoda) a průměr získaných hodnot byl označen za výslednou míru koheze.

Druhým způsobem bylo srovnávání naměřených očních pohybů s trajektoriemi předpovězenými různými modely. Zde se sestavil časoprostor z celé skupiny porovnávaných očních pohybů a spočítaly se hodnoty z buněk odpovídajících předpovědím modelu. V dalších studiích jsme testovali různé scénáře porovnávání (Děchtěrenko et al., 2017) a začali používat korelace místo Normalized Scanpath Saliency. V případě korelací je postup velmi podobný, jen počítáme přímo korelaci dvou časoprostorů. Výhodou je, že korelace nabývá hodnot z menšího rozsahu a výzkumníci mají s těmito hodnotami zkušenosti.

3.1.2 Výsledky

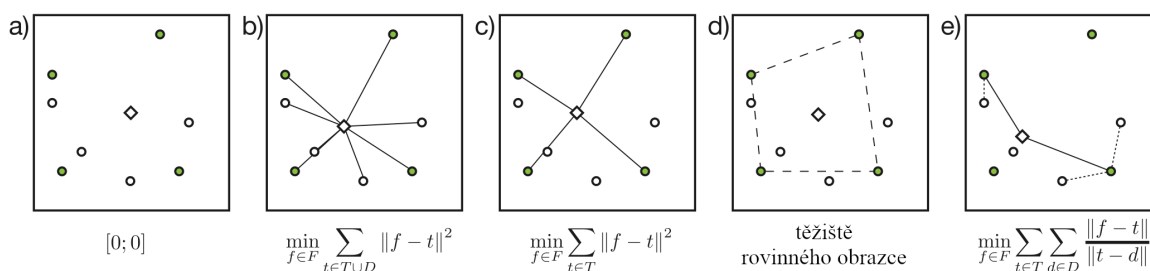
Výsledky studie ukázaly, že lidé oční pohyby při MOT opakují. Při porovnávání podobnosti jsem sledoval vliv dvou faktorů — vliv úlohy (stejná úloha/různé úlohy) a vliv pozorovatele (oční pohyby stejného člověka/různých lidí). Nejvyšší podobnost odpovídala situaci, kdy jsem porovnával oční pohyby ve stejné úloze u stejného člověka. U různých lidí vyvolávala stejná úloha také velmi podobné oční pohyby, i když míra podobnosti byla nižší než v prvním případě. Studie ukázala, že nemá smysl mluvit o individuálním vzorci očních pohybů daného člověka. Oční pohyby stejného člověka v různých úlohách si byly velmi málo podobné a statisticky se nelišily od základní úrovně podobnosti, kterou představovalo porovnávání očních pohybů různých lidí v různých úlohách. Tím, že všechny úlohy se prezentují na monitoru a pohled do blízkosti středu monitoru je poměrně informativní, najdeme nízkou nenulovou základní podobnost i pro nesouvisející oční pohyby.

Přestože se oční pohyby při MOT opakují, opakují se s určitou mírou nepřesnosti či šumu. Pokud stejný člověk sleduje pohybující se objekty na monitoru v zorném poli $20 \times 20^\circ$, oční pohyby opakuje s průměrnou odchylkou 2.5° (3.1° pokud porovnáваме oční pohyby různých lidí). Jen část této odchylky se dá připsat technické chybě měření (dle výrobce v průměru 0.5°). Zbývající část spíše vypovídá o efektivitě plánování očních pohybů v dané úloze — pro určení polohy jednotlivých kroužků lidem tato přesnost stačí a nepotřebují investovat úsilí do přesnějšího cílení očních pohybů. Toto zjištění komplikuje klasifikaci fixací (pohled na cíl, distraktor apod.), jak byla používána v předchozích studiích (Fehd & Seiffert, 2010; Huff et al., 2010; Zelinsky & Neider, 2008). Rovněž tak získáváme horní odhad toho, jak modely očních pohybů mohou být úspěšné a jakou část variability už nebude možné vysvětlit povahou úlohy.

Tato měření nám dávají informaci, v jakém rozmezí se může podobnost očních pohybů pohybovat. Dolním odhadem je zmiňovaná základní úroveň podobnosti (podobnost u různých lidí v různých úlohách), která odpovídá situaci, kdy bychom předvíдали oční pohyby, aniž bychom věděli, kdo je pozorovatelem a jakou MOT úlohu vidí. Horním odhadem je podobnost u stejného člověka ve stejné úloze, která odpovídá situaci, kdy bychom předvíдали oční pohyby známého pozorovatele z jeho vlastních pohybů při předchozích prezentacích dané úlohy. Alternativním horním odhadem může být podobnost ve stejné úloze u různých lidí, protože lze předpokládat, že při předvídání očních pohybů nebude mít model informace o konkrétním pozorovateli a jeho předchozím chování.

Pro zhodnocení současných modelů je zajímavé porovnat, kde se v tomto rozmezí nachází jejich předpovědi. Vybral jsem proto několik modelů založených na počítání středu sledovaných objektů (*all-points centroid* - těžiště všech objektů vč. distraktorů, *target centroid* - těžiště cílových objektů, *object centroid* - těžiště rovinného obrazce definovaného sledovanými body). K modelům jsem přidal triviální model (*constant strategy*), předpokládající, že pozorovatel neuhýbá pohledem a sleduje jen střed obrazovky. Nakonec jsem navrhnul složitější model, který dával vyšší váhu objektům, které jsou v blízkosti jiných objektů a hrozí tak jejich záměna (*anticrowding*). Přestože tento model nebyl do té doby uvažován, odpovídá pozorováním z lokalizačních studií (Iordanescu et al., 2009) a pozorováním tzv. rescue sakád (Zelinsky & Todor, 2010).

Srovnání ukazuje, že úspěšnost tří modelů je velmi podobná (object-centroid, anticrowding, target-centroid). Pokud si rozsah možné podobnosti přepočítáme na škálu 0 až 100 %, dosahují tyto modely 48-54 %. Pokud



Obrázek 3.2: Testované modely očních pohybů u MOT úlohy (Lukavský, 2013). (a) konstantní strategie, (b) těžiště všech objektů (*all-points centroid*), (c) těžiště cílových objektů (*target centroid*), (d) těžiště rovinného obrazce (*object centroid*), (e) strategie minimalizující crowding (*anticrowding*). Cíle jsou zobrazeny zeleně, distraktory bíle, čtvereček značí předpověď modelu.

bychom vycházeli z toho, že nemůžeme mít informace o konkrétním pozorovateli a maximum si definovali z podobnosti s ostatními lidmi, tak modely dosahují úspěšnosti 72-79 %.

Tato čísla se bez kontextu obtížně interpretují a spíše ukazují, že existuje prostor pro další zlepšení modelů. Důležité ale je, že pro MOT jsme vůbec schopni něco takového definovat. V mnoha dalších úlohách jako je prohlížení fotografií či sledování filmu, nemůžeme měřit, jak moc přesně lidé své oční pohyby opakují. To je škoda, protože pak chybí horní odhad toho, jak úspěšné naše modely v těchto úlohách mohou být.

3.2 Statistické testování podobnosti očních pohybů

Předchozí studie ukázala, že lidé v opakovaných situacích provádí sice podobné oční pohyby, ne ale úplně stejné. Povaha úlohy nevyžaduje vysokou přesnost zaměřování fixací a tak při opakované prezentaci vidíme odchylky 2-3°. Pro další výzkum by bylo užitečné mít nástroj, jak ověřit, zda pozorované rozdíly představují běžnou míru variability (jakou vidíme u opakování), nebo zda se jedná o významný rozdíl. V naší další studii (**Děchtěrenko et al., 2017**) jsme se zaměřili na porovnání různých přístupů testování statisticky významných rozdílů u očních pohybů.

V článku rozlišujeme mezi srovnávacími metrikami a srovnávacími strategiemi. Metrikami se myslí metody pro kvantifikaci podobnosti dvou záznamů očních pohybů nebo skupin očních pohybů a výsledkem je číslo. Příkladem mohou být metriky Normalized Scanpath Saliency nebo Correlation Distance zmiňované v předchozí studii. Strategie je statistický způsob, jak porovnat pozorovaný rozdíl mezi očními pohyby s přihlédnutím k obvyklé míře šumu. Má povahu statistického testu a výsledkem tak může být závěr, zda se skupiny očních pohybů od sebe liší nebo ne. Konkrétní strategie stručně představím v podkapitole Metodický přístup.

Náš přístup demonstrujeme na příkladu opakovaného MOT a geometrických manipulací. Jak jsme si již řekli, lidé nedokáží účinně rozlišit, zda se MOT úlohy opakují. Pokud bychom chtěli opakování ještě více zamaskovat, jednou možností by bylo při opakování úlohu zrcadlově překloupit podle vertikální či horizontální osy. Přestože neexistuje silný předpoklad, že by pravolevá symetrie měla výrazně ovlivnit získaná data,³ je regulérní

³Některé studie uvádějí, že lidé mají mírnou preferenci směřovat první pohledy do levé poloviny fotografie (např. Dickinson & Intraub,

se ptát, zda se oční pohyby u původních a zrcadlově překlopených úloh liší. Jak jsme si již řekli, oční pohyby při opakované prezentaci úlohy jsou sice podobné, ale ne úplně stejné. Cílem je tedy navrhnout způsob, jak ověřovat, zda se míra podobnosti v rámci stejné experimentální podmínky (originální vs. originální; příp. překlopená vs. překlopená) liší od podobnosti napříč experimentálními podmínkami (originální vs. překlopená).

3.2.1 Metodický přístup

Navrhli jsme tři strategie pro statistické porovnávání podobnosti očních pohybů. Citlivost těchto přístupů jsme nejprve ověřili na simulovaných datech. V dalším kroku jsme porovnali vliv zrcadlového převrácení úlohy na podobnost očních pohybů (Exp.2 - symetrie podle svislé osy, dále LR-symetrie; Exp. 3 - symetrie podle vodorovné osy, UD-symetrie). U převrácených úloh jsme vždy záznam očních pohybů převrátili zpět tak, aby byl srovnatelný se záznamy z původní úlohy.

Pro všechny strategie předpokládáme, že porovnáváme oční pohyby ve dvou podmínkách (A a B) a v rámci každé podmínky máme několik záznamů (řekněme 6), které nám umožňují odhadnout variabilitu v rámci úlohy.

První navrženou strategií je porovnávání podskupin (*subset comparison*). Strategie spočívá v porovnávání vnitřní konsistence (vzájemné podobnosti) podskupin, které vybíráme buď v rámci stejné podmínky (4 záznamy z podmínky A) nebo napříč podmínkami (2 záznamy A a 2 záznamy B). Záměrně vybíráme podskupiny o menším počtu než kolik je dostupných dat (4 z 6), aby existovalo více možností, jak podskupiny tvořit a získali jsme větší počet dat pro odhad variability. Údaje o konsistenci na závěr porovnáme statistickým testem (např. nepárový t-test).

Druhá strategie párového srovnávání (*pairwise comparison*) vychází z permutačního přístupu (Feusner & Lukoff, 2008). Celkovou vzdálenost mezi podmínkami d^* můžeme vypočítat tak, že spočítáme průměr všech párových vzdáleností napříč podmínkami (1 záznam A vs. 1 záznam B) a odečteme od něj průměr všech párových vzdáleností v rámci podmínek (A vs. A nebo B vs. B). Abychom zjistili, zda je tato vzdálenost d^* výrazně odlišná, spojíme všechny záznamy dohromady (12 záznamů) a budeme vytvářet nové náhodné rozdělení do skupin (6 a 6). U těchto nových rozdělení také spočítáme rozdíl mezi vzdálenostmi napříč skupinami a uvnitř skupin. Až budeme mít dostatek údajů, porovnáme původní hodnotu d^* s celkovým rozdělením a pokud hodnota bude dostatečně odlehlá (např. přesáhne 95. percentil), usoudíme, že se podmínky statisticky významně liší.

Poslední strategie skupinového srovnávání (*groupwise comparison*) kombinuje obě předchozí strategie. Jedná se opět o strategii založenou na permutačním testu, ale základem je porovnávání celých skupin. Nejprve porovnáme skupinu všech záznamů z první podmínky se skupinou záznamů z druhé podmínky (6 záznamů A vs. 6 záznamů B). Nyní záznamy opět sloučíme a budeme vytvářet nová náhodná rozdělení na dvě podmínky. Získáme tak distribuci možných vzdáleností mezi skupinami a budeme porovnávat, jak odlehlá je původní hodnota.

2009). Pokud bychom u MOT předpokládali, že lidé cíle sledují jako skupinu, spíše bychom předpokládali, že se lidé budou dívat do středu skupiny, podobně jako svůj pohled obvykle směřují do středu objektů (Foulsham & Kingstone, 2013).

3.2.2 Výsledky

Simulační část studie ukázala, že všechny tři strategie jsou přibližně stejně citlivé a rozdíly mezi nimi jsou poměrně malé. Pro další experimenty jsme vybrali strategii porovnávání podskupin, protože byla v obou podmínkách simulace citlivější než zbývající dvě strategie.

V dalších dvou experimentech jsme pomocí vybrané strategie testovali vliv symetrického překlopení úlohy. Ukázalo se, že jak převrácení podle svislé, tak převrácení podle vodorovné osy se projeví na očních pohybech. Rozdíl není velký, ale je statisticky detekovatelný. Pokud bychom tedy chtěli zrcadlově převrácené úlohy použít k dalšímu utajení opakování (např. abychom získali větší počet dat), je to možné, ale za cenu zvýšení šumu v experimentu.

Podle předpokladů byly zjištěny větší rozdíly u překlopení podle vodorovné osy (UD-symetrie). Předchozí studie ukázaly (např. Hagenbeek & Van Strien, 2002), že lidé mají odlišnou citlivost v horní a dolní polovině zrakového pole, což může vést k odlišným strategiím, pokud obsah hodní a dolní poloviny prohodíme.

V praxi může volba strategie záviset na vlastnostech použité metriky. Obecně je výběr metriky a strategie nezávislý a je možné stejnou strategii používat s různými metrikami. Nicméně každá ze strategií předpokládá nějakou vlastnost metriky. Porovnávání podskupin předpokládá, že je možné počítat konsistenci (vnitřní podobnost jedné skupiny). Skupinové srovnávání předpokládá možnost porovnávat dvě skupiny proti sobě. Strategie párového srovnávání spoléhá na to, že je možné (a dostatečně robustní) porovnávat mezi sebou jednotlivé záznamy. Sice je možné z jednoho přístupu přecházet na další (např. počítat konsistenci pomocí leave-one-out metody), ale je pravděpodobné, že podle povahy výpočtu některé metriky budou některým strategiím odpovídat lépe než jiným.

Možnost používat různé metriky komplikuje porovnávání výsledků. V případě statických úloh existují srovnání jednotlivých metrik (Dewhurst et al., 2012; Jarodzka et al., 2010), ale pro dynamické úlohy (jako je MOT) podobné srovnání chybělo.

3.3 Oční pohyby a předvídání pohybu

Pokud se snažíme pochopit, kam se lidé dívají při plnění určité úlohy, a chceme toto chování modelovat, porovnáváme aktuální chování člověka (oční pohyby) se situací v prostředí. Důležitou otázkou ale je, zda aktuální stav je pro vysvětlení chování nejdůležitější. Možná by předchozí stavy nebo budoucí stavy prostředí byly relevantnější. Pro obě alternativy existují dobré důvody. Vnímání a plánování určitou dobu trvá, a proto je možné, že chování odráží stav prostředí se zpožděním desítek až stovek milisekund.⁴ Naproti tomu lze argumentovat, že člověk nemusí být pasivní pozorovatel, ale často situaci rozumí (chápe fyzikální zákonitosti, má intuitivní představu o chování aktérů) a dokáže ji předvídat. V tom případě by jeho chování lépe odpovídalo budoucímu stavu prostředí.

Sledování více objektů v úloze MOT je pro účastníky přehlednou a srozumitelnou úlohou. Objekty jsou na

⁴Podle odhadů Thorpe et al. (1996) zpracování obsahu fotografie trvá přibližně 150 ms.

obrazovce zřetelné, pohybují se spojitě a po dobu trvání úlohy je úkol jasný. Právě spojitý pohyb poskytuje dobré podmínky pro možnost předvídání situace. Experimenty kombinující spojitý pohyb a náhlé události (flash-lag effect, Nijhawan, 1994, 2008) ukazují, že se při sledování spojitého pohybu snažíme zpoždění našeho vnímání kompenzovat, a naše subjektivní zkušenost polohy objektu vzniká extrapolací na základě předchozího pohybu.⁵

Ve výzkumu MOT probíhá debata, do jaké míry jsou lidé citliví na informaci o směru objektů. Část vědců argumentovala, že pro sledování používáme pouze informace o poloze objektů (*location-only account*), zatímco jiní poukazovali na to, že lidé informace o směru objektů zpracovávají a používají pro plnění úlohy (*location and direction account*). Argumenty v této debatě pocházely z řady různých experimentálních manipulací. Zkoumalo se například, do jaké míry dokáží lidé kompenzovat pohyb objektů za překážkou (Franconeri et al., 2012; Scholl & Pylyshyn, 1999). V jiných experimentech všechny objekty nečekaně na okamžik zmizeli a testovalo se, zda bude lidský výkon lepší, pokud se objekty navrátí na místa svého zmizení nebo na místa, kam by dojeli, pokud by neměnily směr (Fencsik et al., 2007; Keane & Pylyshyn, 2006). Další experimenty explicitně testovali, zda si lidé pamatují směr objektů (Horowitz & Cohen, 2010), nebo zda je zapamatovaná poloha objektů ovlivněna směrem jejich pohybu (Howard & Holcombe, 2008; Iordanescu et al., 2009).

Howard et al. (2011) navrhují, že výsledné předvídání nebo opoždění je výsledkem několika protichůdných procesů. Opoždění může být způsobeno potřebou integrace podnětů (detekovat pohyb a jeho směr vyžaduje určitý čas). Pokud lidé v nějaké fázi zpracovávají informace sériově (jeden objekt po druhém) nebo pokud ukládání informací vyžaduje hodně kognitivních zdrojů, může to také vést k opoždění. Vliv těchto faktorů mohou lidé kompenzovat pomocí extrapolace pohybu. Náhlé změny, jako je zmizení či objevení objektu, vysílají výrazný signál o poloze objektu, což může narušovat zpracování a výsledky tak nemusí být reprezentativní pro obvyklý průběh úlohy (jako vidíme u flash-lag efektu).

V popisovaném výzkumu jsme si kladli otázku, do jaké míry lidé používají informace o směru pohybu při sledování celé pohyblivé scény. Případné předvídání nebo opoždění očních pohybů za obsahem scény by bylo důležitým faktorem, který by modely očních pohybů měly zohledňovat.

3.3.1 Metodický přístup

Naším cílem bylo prezentovat MOT úlohu bez dalších úprav, aby získaná data bylo možné použít pro modelování obvyklých situací s pohybujícími se objekty. Hlavní myšlenkou bylo, že můžeme lidem ukázat stejnou scénu opakovaně — jednou popředu a později pozpátku. To se dá velmi dobře provést právě u MOT, kde se význam situací při přehrání pozpátku nezmění; naopak u videa s reálnými objekty by tato manipulace přitahovala pozornost. V rámci analýz jsme úlohy, které byly prezentovány „popředu“, opět otočili včetně zaznamenaných očních pohybů. Nakonec jsme vzali související nahrávky „popředu“ a „pozpátku“ a porovnávali pomocí metod popsanych dříve (Děchtěrenko et al., 2017; Lukavský, 2013). Při porovnávání jsme zkoušeli různé posuny v čase a hledali jsme nejvyšší míru shody.

⁵V experimentech je zpoždění názorně viditelné. Pokud doprostřed pohybujícího se prstence náhle umístíme menší kontrastní kruh, lidé vidí kruh posunutý. Zatímco poloha pohybujícího se prstence je extrapolovaná, prezentace kruhu se nedala předvídat (Nijhawan & Khurana, 2000).

Pro ilustraci si představme, že oční pohyby lidí jsou opožděné za sledovaným obsahem o 100 ms. U nahrávky „pozpátku“ budou (po otočení) oční pohyby naopak předcházet sledovaný obsah o 100 ms. Když vezmeme obě nahrávky a budeme je vzájemně posunovat v čase a hledat nejlepší shodu, měli bychom ji najít pro posun o $2 \times 100 = 200$ ms.

Tímto způsobem jsme testovali obecnou úroveň předvídání/opoždování (Exp. 1). V dalších experimentech jsme měřili, jaký vliv má větší kognitivní náročnost úlohy (Exp. 2) a předvídatelnost pohybu objektů (Exp. 3a a 3b).

3.3.2 Výsledky

Naše výsledky jednoznačně ukazují, že oční pohyby se za obsahem scény opoždují. Toto zpoždění činí v průměru asi 110 ms. Mezi lidmi jsme našli rozdíly, nicméně u všech účastníků bylo zpoždění patrné.⁶ Další experimenty ukázaly, že je možné experimentální manipulací zpoždění dále zvýšit, pokud se budou objekty pohybovat více chaoticky (asi o 30 ms). Naopak manipulace, které by potenciálně činily úlohu jednodušší a usnadňovaly předvídání, významnou změnu nepřinesly. Sice byly pozorováno menší zpoždění (o 15 ms u sledování menšího počtu objektů; o 19 ms u více přímých pohybů), ale v rámci pozorované variability tyto výsledky nejsou statisticky významné.

Z našich výsledků plyne, že pokud hledáme strategie, kam se lidé dívají při sledování pohybujících se objektů, neměli bychom aktuální polohu očí srovnávat s aktuálními pozicemi objektů v prostředí, ale s jejich předchozími pozicemi (zhruba o 110 ms zpět). Výsledky nevyvrací možnost, že lidé předvídají směr pohybu objektů. Pokud se na situaci budeme dívat jako na souběh více protichůdných faktorů (Howard et al., 2011), k extrapolaci může stále docházet jen nemá tak výrazný efekt. Extrapolace může mít význam pro zaměření pozornosti, které se ale nebude projevovat v očních pohybech. Lidé tak mohou sledovat objekty se zpožděním, ale stále být citliví na drobná bliknutí ve směru pohybu objektů (Atsma et al., 2012). Tato disociace mezi směrem pohybu a zaměřením pozornosti může souviset s jejich odlišnou mírou složitosti a potřebnými náklady. Zatímco změna zaměření pozornosti vyžaduje pouze kognitivní úsilí, pro pohyby očí je třeba zapojit okohybné svaly.

3.4 Pozornost a vliv distraktorů

Předchozí studie se zabývaly dynamickými úlohami, kde byla důležitá poloha objektů, ale objekty samy byly velmi jednoduché. Předložená studie (Lukavský, 2019) zkoumá otázku, do jaké míry dokážeme směřovat pozornost na část zorného pole a ignorovat rušivé informace ze zbytku pole. Podobně jako v MOT úlohách předem víme, které části zorného pole nás mají zajímat a které máme ignorovat. Rozdělení na objekty, ale již nebude tak přímočaré jako u MOT (4 cíle a 4 distraktory). Tentokrát budeme lidem ukazovat části fotografií (ulice, les, nádraží). Každá část obsahuje různé objekty v přirozených polohách. Úkolem účastníka bude se v krátkém čase ve fotografii vyznat a kategorizovat její obsah (zda se jedná o přírodní nebo lidmi vytvořenou scénu), současně se ale nesmí nechat zmást jinou fotografií zobrazenou ve zbývající části zorného pole.

⁶Průměrná zpoždění se pohybovala v rozmezí mezi 50 a 170 ms. Náš design nebyl navržen pro měření individuálních rozdílů, proto ze zjištěných individuálních výkonů nic dalšího nevysuzujeme.

Na obecné úrovni se tato úloha podobá například známé Stroopově úloze (Stroop, 1935), kde lidé vidí v kritické podmínce slova označující barvy napsaná různými barvami (slovo „červený“ napsané modře). Jejich úkolem je určit barvu písma a nenechat se zmást obsahem textu. Ukazuje se, že je to pro lidi poměrně obtížné a automaticky text čtou, i když je pro úlohu nejen nerelevantní, ale dokonce matoucí. Podobný vliv psaného textu najdeme i v lingvistických výzkumech (Starreveld & Heij, 2017), kde lidé vidí současně napsané slovo a obrázek předmětu. Mají za úkol buď přečíst slovo nebo pojmenovat obrázek (picture-word interference paradigm). I zde se ukazuje, že lidé text automaticky čtou a vliv čtení je výraznější než vliv zpracování obrázku (Rosinski et al., 1975). Pokud úkol vyžaduje hlubší zpracování můžeme vliv zpracování obrázku pozorovat (Greene & Fei-Fei, 2014). V úlohách zmíněných výše účastníci sledovali dva kanály informací, jejichž obsah mohl být ve vzájemném konfliktu. Je potřeba zdůraznit, že oba informační kanály nebyly rovnocenné, vyžadovaly různé způsoby zpracování (čtení a rozpoznávání objektů), a tak role pozornosti je zde spojená spíše s koordinací obou činností než se směřováním na různá místa zorného pole.

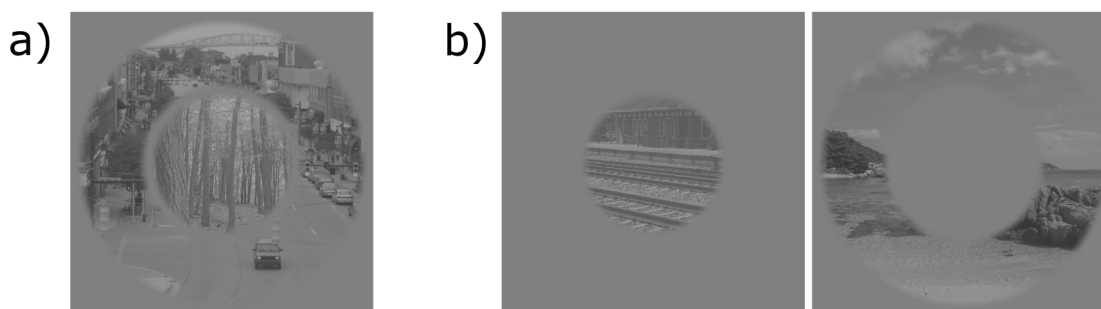
Některé studie navozují konflikt v rámci zorného pole při použití podobných způsobů zpracování. Navon (1977) ve známé studii prezentoval obrazy písmen složených z jiných písmen („H“ složené z menších „S“). Jeho výsledky ukazují, že lidé zpracovávají globální tvar obrazu rychleji než detaily. Naopak výzkumy fotografií s nekonzistentním obsahem (např. fotbalista v katedrále) spíše dokládají symetrický vztah mezi figurou a pozadím. Tedy že v případě konfliktu mezi dominantní postavou scény a kontextem, je narušeno zpracování jak postavy, tak kontextu (Davenport, 2007; Davenport & Potter, 2004).

Aktuální studie si klade otázku, jak lidé zpracovávají informace z odlišných míst, které mohou být v sémantickém konfliktu (viz Obr. 3.3). Celkem jsme navrhli a experimentálně testovali tři možné hypotézy:

- 1) Zpracování je izolované a nezávislé; v případě konfliktu budou lidé schopni potlačit nesouvisející signál. To by jim mohla umožňovat anatomická oddělení oblastí mozku, které informace z různých částí zorného pole zpracovávají, případně lidská schopnost rozdělit vizuální pozornost do několika odlišných míst zorného pole (Franconeri et al., 2007). I strategické rozhodnutí, zda se lidé rozhodnou spoléhat na detaily nebo celkové parametry, může pomoci izolovat oba informační kanály.
- 2) Zpracování je automatické s prioritou pro jeden kanál, podobně jako to vidíme u Stroopova efektu. V tom případě bychom čekali, že rozhodování o prioritním kanálu nebude ovlivněno obsahem druhého, ale ne naopak. Např. hypotéza rozšiřující se pozornosti (zoom-out hypothesis, Larson et al., 2014) předpokládá, že při rychlém zpracování obrazu dostane prioritu střed zorného pole. Naproti tomu představa, že vizuální zpracování začíná hrubými obsahy a postupuje k detailům (coarse-to-fine recognition, e.g., Navon, 1977; Schyns & Oliva, 1994), by předpovídala, že informace z periferie by mohly být dostupné dříve.
- 3) Zpracování je současné a vzájemně se ovlivňuje; v případě konfliktu uvidíme zhoršení u obou kanálů, podobně jako u studií s nekonzistentní figurou a pozadím (Davenport, 2007; Davenport & Potter, 2004).

3.4.1 Metodický přístup

Základem experimentu byla prezentace kombinovaných fotografií, které se skládají ze dvou částí: centrálního kruhu a širokého prstence v jeho okolí (viz Obr. 3.3a). Každá část obsahovala jinou fotografii. Účastník byl vždy předem upozorněn, zda se má soustředit na střed nebo okolní prsteneček. Abychom zamezili tomu, že lidé nebudou vědět, na co mají dávat pozor, cíl byl vždy stejný pro delší skupiny úloh (měnil se po 50 úlohách). V každé úloze bylo úkolem určit, zda fotografie v cílové oblasti znázorňuje přírodní nebo lidmi vytvořenou scénu. V obou případech mohly fotografie zachycovat různé kategorie scén (přírodní scény zahrnovaly hory, les, pláže, pouště nebo řeku; lidmi vytvořené scény domy, bazény, tržiště, ulice a nádraží). Fotografie byly černobílé, aby se nedalo spoléhat na typické barvy (Oliva & Schyns, 2000).



Obrázek 3.3: Ukázka podnětů. (a) Kompozitní fotografie používaná při experimentu (Lukavský, 2019). (b) Fotografie zachycující pouze střed a periferii, používané v pilotní studii. Podobné fotografie byly používané v předchozích studiích (Larson et al., 2014; Larson & Loschky, 2009).

Velikost fotografií vycházela z předchozích studií (Larson et al., 2014; Larson & Loschky, 2009), které se zabývaly schopností lidí určit povahu scény ze středu a periferie zorného pole. Obě oblasti přispívají k rozpoznání vlastním způsobem. Signály ze středu jsou ostřejší, zachycují více detailů a umožňují lepší identifikaci objektů, které mohou být pro rozpoznání důležité (Oliva & Schyns, 1997). Lidé tak například mohou snáze poznat lidmi vytvořenou scénu podle přítomnosti aut. Signál z periferního vidění neobsahuje takové detaily, ale je zpracováván poměrně rychle a obsahuje dostatečné informace pro rozpoznávání scén (gist, Torralba & Oliva, 2003). Larson a Loschky (2009) hledali, v jakém poměru musí být střed a okolí, aby lidé byli při rozpoznávání obsahu stejně úspěšní. V pozdější studii (Larson et al., 2014) manipulovali dobu prezentace a testovali, který obsah je zpracováván dříve. To je vedlo k formulaci hypotézy o rozšiřující se pozornosti (zoom-out hypothesis).

Je dobré předeslat, že použité fotografie byly vybrány tak, aby jejich klasifikace nebyla rozporuplná. Pokud by lidé měli dostatek času, na zařazení fotografií by se shodli. Aby bylo možné zjistit, jak úspěšně zpracování probíhá v jednotlivých podmínkách, úloha byla ztížena nedostatkem času. Fotografie byla zobrazena jen na okamžik (200 ms) a vzápětí nahrazena vhodně zvolenou maskou⁷. Zejména při rozhodování o obsahu periferní části je přirozené, že by tam lidé měli tendenci přesunout pohled. Proto bylo na začátku každé úlohy kontrolo-

⁷Masky se vytváří tak, aby zachovaly některé základní spektrální charakteristiky původní fotografie, ale neobsahovaly žádné rozpoznatelné informace (Loschky et al., 2010; Portilla & Simoncelli, 2000).

váno pomocí eye-trackeru, zda se lidé skutečně dívají do středu fotografie. Krátká doba prezentace odpovídá přibližně době jedné fixace, a proto v průběhu prezentace již lidé nestíhají přesunout svůj pohled na okraj.

3.4.2 Výsledky

Obsahem studie jsou dva experimenty. V obou experimentech se sleduje úspěšnost rozpoznání v závislosti na dvou klíčových faktorech (2×2 design). Prvním faktorem je místo zaměření pozornosti, tedy zda se měl účastník zaměřit na rozpoznání centrální nebo periferní části obrazu. Druhým faktorem je přítomnost konfliktu mezi obsahy obou prezentovaných částí; jinými slovy zda obě patřily do stejné kategorie (přírodní/lidmi vytvořená scéna) nebo ne. Druhý experiment ještě přidával třetí faktor: čas na zpracování. Cílem bylo zjistit, zda se výsledky mění v závislosti na množství času, které bude mít účastník k rozpoznání obrazu.⁸

Oba experimenty přinesly podobné výsledky. Zaprvé, lidé byli úspěšnější při rozpoznávání centrální části. Zadruhé, přítomnost konfliktu zhoršovala výkon pro obě části, centrální i periferní. To je klíčový výsledek, protože je v rozporu s předpoklady zoom-out hypotézy. Pokud se vrátíme k hypotézám popsaným výše, tak tento výsledek podporuje představu o současném a vzájemně se ovlivňujícím zpracování sledovaných i nesledovaných částí obrazu. I když bylo zpracování obsahu centrální části úspěšnější, nejedná se o prioritní zpracování (jako u Stroopova efektu), protože konfliktní signály z periferie, na které se účastníci neměli soustředit, zhoršovaly výkon.

Teoreticky by bylo možné, že by se prioritní zpracování centrálního signálu projevovalo pouze v počátečních fázích a při 200 ms již nenacházíme rozdíl. Tedy slovy zoom-out hypotézy, že se pozornost rozšiřuje z původního centrálního zaměření a v daném okamžiku již pokrývá celou oblast s oběma podněty. Experiment 2 však ukázal, že periferní signál ovlivňuje zaměřené zpracování centrálního signálu již při velmi krátké prezentaci (33 ms).

Tento vliv podnětů, které měly být ignorovány, můžeme vidět i v dynamických úlohách s rozdělenou pozorností. V jedné předchozí studii (Děchtěrenko & Lukavský, 2016) jsme porovnávali pomocí strojového učení, které údaje o objektech vedou k nejpřesnějšímu předvídání očních pohybů při MOT. Předvídání očních pohybů je relativně přesné, (1) pokud víme, kam se člověk díval při předchozí prezentaci této úlohy (úloha je opakována), nebo (2) pokud víme, kam se dívali v předchozím okamžiku videa (pohled se většinou při MOT posunuje spojitě). Pokud tyto informace nemáme, jsme odkázáni pouze na geometrické informace o poloze a vzájemných vzdálenostech jednotlivých objektů, tedy cílů a distraktorů. Většina teorií předpokládá, že lidé sledují cíle a změny u distraktorů mají malý vliv, i když experimentální data ukazují, že charakteristiky distraktorů vliv mají. Například výkon lidí je lepší, pokud se MOT úloha zopakuje celá oproti situaci, kdy se opakuje pouze pohyb cílů (Ogawa et al., 2009). Proměnlivá rychlost objektů činí úlohu obtížnější i v situacích, kdy takto měníme pouze rychlosti distraktorů (Meyerhoff et al., 2016). Lidé polohu distraktorů částečně reprezentují (Alvarez & Oliva, 2008) a náhlé změny polohy distraktorů zhoršují výkon v úloze, i když jsou zachovány vzdálenosti k cílům (Meyerhoff et al., 2015).

Naše srovnání ukázalo, že nejpřesnější predikce skutečně dosahují modely, které zohledňují polohu distraktorů. Z teoretických modelů byla nejpřesnější anti-crowding strategie (Lukavský, 2013), která předpokládá, že váhu

⁸V Experimentu 1 viděli účastníci fotografie po dobu 200 ms. V Experimentu 2 byl zkoumán vliv kratší i delší doby (33, 200 a 400 ms).

jednotlivým cílům přisuzujeme podle jejich blízkosti k dalším objektům. Z empirických modelů založených na strojovém učení, měly výhodu modely, které znaly polohu všech objektů, zejména pak i model, který pracoval se vzdálenostmi mezi cíli a distraktory. Jak již bylo zmíněno, problémem dalšího vývoje teorií MOT jsou velmi podobné predikce stávajících teorií. Naše studie (Děchtěrenko & Lukavský, 2016) je cenná v tom, že ukazuje, že pomocí strojového učení můžeme porovnávat, které informace jsou pro zpracování důležité a pomáhají replikovat lidské chování, a které ne. Budoucí teorie pozornosti při MOT by měly uvažovat nejen zaměření a distribuci pozornosti na sledované cíle, ale také inhibiční vlastnosti pozornosti a jak s nimi potlačujeme rušivé vlivy během úlohy. Podobné úvahy již vidíme v nových teoriích jako je interferenční teorie (Franconeri et al., 2010).

Kapitola 4

Shrnutí metodologického přínosu

Současné sledování různých objektů v našem okolí je častý úkol, který naše mysl řeší. Čtyři diskutované studie systematicky zkoumaly¹, jak se sledování objektů projevuje navenek v očních pohybech. Přestože součástí studií je i vlastní návrh modelu (anticrowding), hlavní přínos vidím v metodologických aspektech mapování faktorů, které oční pohyby ovlivňují. Práce jsou zahrnuty v současných přehledových studiích o MOT (Meyerhoff et al., 2017) i o očních pohybech během MOT (Hyönä et al., 2019). Také jsou citovány autory nových modelů (Upadhyayula & Flombaum, 2020) a popisované metody jsou dál rozvíjeny (Salas & Levin, 2021). Celková zjištění napříč jednotlivými studiemi shrnuji v následujících bodech.

- 1) Pokud dokážeme najít úlohu, která se dá smysluplně opakovat, je možné odhadnout množství šumu v našich očních pohybech. V případě Multiple Object Tracking sice lidé své oční pohyby opakují, ale jen přibližně (průměrná odchylka 2.5°).
- 2) Znalost míry šumu nám umožňuje odhadnout, kolik variability dokáží vysvětlit současné modely očních pohybů a kolik variability zůstává nevysvětleno. Současné modely vysvětlují přibližně 75 % variability, pokud nebudeme aspirovat na vysvětlení interindividuálních rozdílů. Na úrovni individuálních výkonů vysvětlují přibližně 50 % variability.
- 3) Přítomnost šumu v měření znamená, že při srovnávání různých experimentálních podmínek potřebujeme pomocí statistických testů vyjádřit, jak se liší míra šumu uvnitř stejné podmínky a napříč podmínkami. Přestože se dá k problému přistupovat různými způsoby, v závislosti na tom, zda porovnáváme dvojice záznamů nebo celé skupiny, výsledky jsou velmi podobné.
- 4) Metody jako Normalized Scanpath Saliency nebo Correlation Distance umožňují přesnější měření očních pohybů v dynamických úlohách. Odpadají problémy s potřebou kategorizace míst, na které se pozorovatel dívá.
- 5) Tři současné modely (*target centroid*, *object centroid*, *anticrowding*) jsou ve svých předpovědích podobně úspěšné.

¹Práce odkazovala i na výsledky dvou nezahrnutých prací publikovaných v zahraničních recenzovaných sbornících (Dechterenko & Lukavsky, 2018; Děchtěrenko & Lukavský, 2016).

- 6) Informace o nerelevantních objektech (distraktorech) jsou pro vysvětlení očních pohybů důležité a zlepšují predikci při strojovém učení. Ze zmíněných tří modelů pouze *anticrowding* model pracuje s polohou distraktorů a tak se jeví perspektivnější než zbývající modely.
- 7) Lidé nedokáží plně ignorovat nerelevantní podněty. I když víme, že obraz na periférii nebude podstatný, bude rušit naše rozhodování o centrálně umístěném podnětu, a to už i v raných fázích zpracování.
- 8) V dynamické úloze je třeba zvažovat, zda oční pohyby souvisí s aktuálním stavem prostředí a zda by nebylo lepší je vysvětlovat předchozími či budoucími pozorováními. Naše pozorování ukázala, že v případě MOT se oční pohyby za obsahem scény opožďují asi o 110 ms. U jednotlivých lidí existují v míře zpoždění drobné rozdíly, ale u všech se jedná o zpoždění alespoň o 50 ms.
- 9) Zpoždění očních pohybů oproti obsahu scény lze prodloužit ztížením predikce pohybu, ale ke zkrácení při dobrých podmínkách nedochází. Ani více kognitivní kapacity díky sledování menšího počtu objektů, ani lépe předvídatelný pohyb nemá vliv.

Jednotlivé práce se převážně zabývají modelovou situací Multiple Object Tracking. Pro další teoretický rozvoj považuji za vhodné se zaměřit na rozšíření tohoto modelu směrem k dalším kognitivním funkcím a zkoumat podrobněji podíl vnímání a pracovní paměti. V tomto modelu předpokládáme, že není problém objekty rozeznat od sebe nebo oproti pozadí. V reálných situacích ale objekty mohou být různě patrné, a tak analýza našich schopností sledovat objekty navzdory vizuálnímu šumu by byla přínosná například pro testování kamufláže a kontrakamufláže. Zajímavé jsou situace, kdy si potřebujeme o objektech pamatovat další informace — který z maskovaných lupičů nese hledaný předmět nebo z které strany čekáme, že se vrátí objekt, který zmizel za okrajem zorného pole. Velkou mezeru v současném výzkumu vidím v nevyužitém diagnostickém potenciálu úlohy, na který bych se chtěl v budoucnu zaměřit.

Kapitola 5

Aplikované analýzy očních pohybů

Hlavní přínosy metodologické části byly prezentovány na dynamických úlohách jako je MOT. Tyto metody rozšiřující možnosti, co se můžeme z dat o očních pohybech dozvědět. Nabízí se otázka, zda a jak konkrétně by se tyto metody daly využít v dalších výzkumech. Rád bych se u tohoto problému zastavil a ukázal potenciál představených postupů.

Za tímto účelem jsem vybral čtyři aplikované studie, na kterých jsem se podílel a které metodu sledování očních pohybů používají. Nejedná se o metodologické studie — jednotlivé studie řeší odborné otázky z pedagogické psychologie nebo problematiku vnímání umění. Jednotlivé studie stručně představím kvůli kontextu, ale hlavně se zaměřím na to, jak studie využívala sledování očních pohybů a jak by nové přístupy mohly pomoci. Nemá se jednat o kritiku jednotlivých studií — tyto studie používají v současné praxi obvyklé postupy a jejich publikace svědčí o tom, že své závěry zdůvodňují přesvědčivě. Cílem je upozornit na omezení obvyklých postupů a na možnosti, jak upravovat design studií, aby bylo možné získat další informace a přesvědčivější doklady pro výzkumné otázky.

Jednotlivé studie se liší ve svém experimentálním designu. Některé porovnávají, jak se liší oční pohyby u různých druhů podnětů s tím, že účastníci vždy viděli buď jen jeden druh (Stárková et al., 2019) a nebo všechny druhy (Adámek et al., 2019; Kesner et al., 2018). Jinde jsou prezentované podněty v podstatě stejné a liší se jen zvýrazněním některých prvků, které jsou pro hlavní úlohu irelevantní (Grygarová et al., 2019). V neposlední řadě jsme zkoumali i mezilidské rozdíly, kdy stejné podněty prohlížely různé skupiny účastníků (Adámek et al., 2019).

5.1 Vliv podoby studijních materiálů na efektivitu učení

První představená studie (Stárková et al., 2019) se zabývala otázkou, jak podoba studijních materiálů ovlivňuje proces a efektivitu učení. Při přípravě studijních materiálů je cílem autorů nejen přehledně strukturovat informace, ale také studenta zaujmout a motivovat k učení. Kognitivní aspekty přípravy materiálů jsou již delší dobu studovány a jsou formulována konkrétní doporučení (např. Mayer, 2009). Pro emoční a motivační aspekty

přípravy zatím detailní a ověřená doporučení chybí. Jednou z možností, jak materiály oživit, je antropomorfizace grafických prvků. Konkrétně si to můžeme představit jako přidání očí či rukou jednotlivým prvkům, aby více působily jako aktéři a komunikovaly dynamiku popisovaného procesu (např. u viru záměr zaútočit).

V popisované studii účastníci měli za úkol vlastním tempem prostudovat přibližně 10 snímků s obrázky a textovými komentáři, které vysvětlovaly průběh nákazy buňky virem chřipky. Materiály byly připraveny v několika verzích, které se lišily mírou antropomorfizace a také použitím barev. Bylo sledováno, kolik si účastníci zapamatují a jak hodnotí svoji náladu během učební situace. Sledováním očních pohybů jsme testovali, jak antropomorfizace ovlivní distribuci pozornosti během učení. Popisovaná studie čerpá z předchozího výzkumu (Mayer & Estrella, 2014), ale významně jej rozšiřuje: používá eye-tracker, nezávisle manipuluje přítomnost barev a zkoumá výsledky učení v delším časovém odstupu.

Samostatným metodologickým problémem, který se sledováním očních pohybů souvisí, byly negativní emoce (zejm. nervozita z přítomnosti eye-trackeru) zmiňované účastníky pilotní studie (Brom et al., 2016). Protože podobné pocity by mohly kontaminovat pocity spojené s antropomorfizací, byl eye-tracker používán jen v polovině měření. To umožňovalo srovnání a v hlavní studii již tato nervozita nebyla replikována.

Při měření očních pohybů jsme si definovali oblasti zájmu (AOI). Na základní úrovni se jednalo o hlavní grafické prvky na studovaných obrázcích (buňka, virus, pole s texty, šipky). Porovnávali jsme, kolik času pohled účastníků směřoval do jednotlivých oblastí. Na obecnější úrovni jsme si objekty rozdělili na grafické a textové prvky a porovnávali dobu strávenou jejich studiem. V analýze jsme se také zaměřili na první 2 vteřiny prezentace každého snímku, abychom zachytili, jaké prvky přitáhnou pozornost jako první. Výsledky ukázaly, že antropomorfizované obrázky přitáhnou více pozornosti v prvních vteřinách oproti kontrolním obrázkům, ale během celé doby studia již tento rozdíl není patrný. Celkově účastníci strávili přibližně 6× víc času studiem textů oproti prohlížení grafických prvků.

Experiment v současné podobě dobře odráží přirozené podmínky při učení: účastníci postupují vlastním tempem a mají srozumitelně zadaný úkol. Analýza přes AOI je snadno interpretovatelná, ale přináší také určitá omezení. Tím hlavní je *potřeba jednoznačného zařazení*. Pro každé místo na obrazovce je potřeba jednoznačně rozhodnout, do které AOI patří. Při čtení textu se dá předpokládat, že lidé budou fixovat pohled v blízkosti jednotlivých slov. U obrázků ale není jasné, zda člověk potřebuje informace tak přesné nebo zda pohled do blízkosti prvků či mezi ně je dostatečný. Takové pohledy pak ale bude obtížně zařazovat. Zde by mohla pomoci spojitá hodnocení podobnosti očních pohybů (pomocí NSS nebo korelací), která byla popsána výše.

Bylo by tak možné souhrnně porovnat podobnost očních pohybů účastníků s antropomorfizovanými podněty oproti skupině s kontrolními podněty. Tato podobnost by se dala testovat souhrnně a nebo pro jednotlivé účastníky *se zohledněním přirozené variability dat*, jak jsme navrhovali v jedné z předchozích studií (Děchtěrenko et al., 2017).

Určitou komplikací by byla proměnlivá délka prezentace snímků, protože každý účastník postupoval různou rychlostí. To je řešitelný problém, jednou z možností je chápat data staticky a sloučit údaje o fixacích z celého snímku. Podobným způsobem ostatně pracuje i použitá analýza AOI. Porovnávání pomocí NSS sice umožňuje srovnání v čase (Dorr et al., 2010; Lukavský, 2013), ale je možné jej používat i pro statické scény (Peters et al.,

2005).

Experiment by bylo zajímavé rozšířit o *dynamickou stránku* a místo statických snímků prezentovat videa či komentované animace. Taková úprava by byla také ekologicky validní (videa jsou častý výukový materiál) a současně by řešila problém proměnlivé délky prezentací. Bylo by pak možné sledovat, jak se liší distribuce pozornosti v určitý moment nebo v závislosti na právě slyšené informaci. Použitá analýza přes AOI by byla pro tyto účely obtížně použitelná resp. příliš pracná, protože jednotlivé oblasti by nestačilo popsat jednou pro celý snímek, ale i jejich pohyb. Naopak spojitá hodnocení pozornosti by bylo možné používat přímo bez dalších příprav podobně jako v úlohách MOT.

5.2 Směr pohledu portrétovaných osob

V druhé studii (Kesner et al., 2018) jsme se zabývali otázkou, jaký vliv na diváka má směr pohledu portrétovaných osob na prohlížených obrazech. Studie měla dvě části, v první jsme sledovali mozkovou aktivitu pomocí fMRI, v druhé jsme samostatně měřili oční pohyby. Během experimentu byly prezentovány portréty z různých časových období od renesance po současnost. Jednalo se o velmi rozmanité obrazy, ale alespoň pro základní standardizaci stylu byly od každého autora zařazeny dva obrazy: na jednom se portrétovaná osoba dívala na diváka, na druhém jiným směrem (direct/averted gaze). Cílem eye-trackingové části studie bylo zjistit, zda přímý pohled povede u diváků k většímu zájmu o oblast očí portrétované osoby.

Účastníci byly požádáni, aby si portréty „prohlíželi jako v galerii“. Portréty byly prezentovány na monitoru počítače a účastníci měli na každý portrét omezený čas 6 vteřin. Během příprav byly pro každý portrét definovány 3 oblasti zájmu (AOI): oblast očí, oblast úst a zbývající části obličeje. U každé oblasti jsme sledovali celkovou dobu fixací. Výsledky ukázaly, že lidé nejčastěji sledují oči portrétovaných osob (přibližně 25 % času) a ústa sledují mnohem méně (5-10 %). Pokud se na ně portrétovaná osoba dívá, věnují sledování jejich očí ještě více času.

Podobně jako v případě předchozí studie by bylo možné aplikovat *spojitá hodnocení podobnosti*, aby se odstranila nutnost každé místo pohledu jednoznačně zařadit. Je totiž možné, že zejména při zmenšení portrétů na plochu monitoru lidé nepotřebují cílit pohled tak přesně do relativně malých AOI. Hlavní výhoda tohoto přístupu by se ukázala v případě, že by každý portrét existoval v obou variantách (jedna s přímým, druhá s odvráceným pohledem). To by vyžadovalo digitální manipulaci původních materiálů, což by nemuselo odpovídat celkové kompozici obrazu a mohlo vyznít nepřirozeně. V případě, že by se použily úplně nové portréty, např. fotografie, jak je i navrhováno v diskuzi článku, mohlo by se s tím již počítat. Potom by bylo možné provést přímé srovnání vlivu směru pohledu na oční pohyby diváka včetně časového srovnání, kdy k rozdílům dochází.

Dalším aspektem, který by mohlo být zajímavé zvážit, je *míra šumu v datech*. Účastníci dostávali totiž poměrně obecnou instrukci a řada jejich očních pohybů nemusela souviset s obsahem obrazu a v jiných situacích by takové oční pohyby již neopakovali. Na rozdíl od MOT by asi nebylo vhodné měřit míru šumu pomocí opakované prezentace podnětů stejným lidem, protože by si portréty pamatovali. Ale pro odhad míry šumu by mohlo sloužit i srovnání napříč jednotlivými diváky. Z dat by bylo možné zjistit, na kterých místech pohledu se

většina lidí shodne a která místa jsou fixována jen malým množstvím diváků či jednotlivci. Dalo by se tak například kvantifikovat, jak se liší oční pohyby konkrétního diváka od průměru. Toto srovnání by se dalo provádět nejen v prostoru, ale i v čase. Dalo by se tak například zjistit, že většina lidí začíná prohlížení portrétu určitými místy a již se k těmto místům nevrací (jejich opakovaná návštěva v průběhu prohlížení je málo pravděpodobná). Jiná místa naopak mohou přitahovat pozornost v průběhu celé doby prohlížení. Podobná měření by mohla být podkladem pro kvalitativní analýzu obrazů, kde by se dalo hledat, jaké významy či umělecké techniky takto působí.

5.3 Vnímání fotografií v kontextu

Ve výzkumných studiích jsou většinou fotografie zkoumány izolovaně. Fotografie jsou prezentovány samostatně a nepředpokládá se, že by na výsledný dojem z aktuální fotografie měly předcházející fotografie vliv (Schupp et al., 2012). V tištěných i internetových médiích fotografie soupeří o naši pozornost a z praktického hlediska je zajímavě vědět, jaký vliv má současné zobrazení různých fotografií. Můžeme se např. ptát, zda zvýrazněním některých prvků převládne celkový pozitivní nebo naopak negativní dojem.

Právě takovou situaci jsme se snažili modelovat ve třetí studii (Grygarová et al., 2019). Účastníci studie si prohlíželi obrazovku obsahující šest emočně nabitých fotografií a dvě tváře. Fotografie byly vybrány z databáze IAPS, která se pro výzkum často používá (Lang et al., 2008) a jejíž součástí je i průměrné hodnocení, jak pozitivně/negativně fotografie na pozorovatele působí. Z šesti vybraných fotografií v každé kompozici byly tři negativní a tři pozitivní. Kompozice byla doprovázena dvěma obrazy tváří (vždy jedna s radostným a jedna se smutným výrazem). Na prohlížení každé kompozice měli účastníci 3 vteřiny, což je poměrně málo, ale můžeme si to představit jako rychlý pohled na stránku časopisu nebo webu, kde bychom se teprve rozhodovali, zda se zajímat o podrobnosti nebo jít dál.

U každé kompozice jsme sledovali, která místa upoutají pohled účastníků (konkrétně kolik času stráví sledováním negativních/pozitivních fotografií) a také jsme se ptali na celkový emoční dojem (na škále 1-6). Protože byly v každé kompozici počty negativních a pozitivních fotografií vyrovnané a i emoce zobrazených tváří byly vyrovnané, není důvod mezi jednotlivými kompozicemi předem očekávat velké rozdíly. V rámci experimentu jsme manipulovali dvěma podmínkami. První podmínkou bylo zvýraznění jedné z tváří příp. absence zvýraznění. Celkově tedy šlo o tři možnosti (zvýrazněna radost/smutek/nic). Druhou podmínkou byla poloha tváří ve vztahu k fotografiím. Obě tváře byly zobrazeny buď uprostřed (blízko místa úvodní fixace) a nebo na okrajích kompozice. Cílem těchto manipulací bylo zjistit, zda zvýraznění určitého emočního prvku bude mít vliv na následné prohlížení fotografií a na celkové emoční hodnocení.

Výsledky ukázaly, že účastníci se skoro nedívali na tváře a většinu času věnovali prohlížení fotografií. To bylo v souladu se záměrem studie, tváře byly zobrazeny jako velké a snadno rozpoznatelné i periferním viděním. Také kromě emočního výrazu nenesly další význam, aby na sebe nestrhávaly pozornost na úkor fotografií. Ukázalo se, že tváře měly velmi malý vliv, pokud byly zobrazeny uprostřed kompozice, nezávisle na tom, jaká emoce byla zvýrazněna. Když byly tváře zobrazeny na okrajích, zvýraznění mělo vliv na to, které fotografie budou účastníci sledovat víc (zvýraznění smutné tváře vedlo ke kratšímu sledování pozitivně laděných fotografií a

naopak). Celkově experiment ukázal, že je možné ovlivnit způsob prohlížení kompozice pomocí emočních vodítek, i když jsme neočekávali, že centrální vodítka budou až tak neúčinná.

Tato studie ilustruje, že i *irrelevantní podněty* ovlivňují lidské chování a způsob prozkoumávání. To je v souladu s našimi předchozími zjištěními, kdy periferní podněty ovlivňovaly hodnocení podnětů zobrazených ve středu zorného pole (Lukavský, 2019). Z metodologického hlediska je důležité, že se nemůžeme spoléhat na to, že vše podstatné vyčteme ze samotných dat (data-driven analýza). Vliv nefixovaných podnětů v periferním poli ukazuje, že pro analýzy je potřeba zahrnovat i další informace o experimentální situaci, pokud je máme k dispozici.

Oproti předchozím studiím, kde by *spojitá hodnocení podobnosti* očních pohybů byla přínosem, zde by benefit nebyl tak výrazný, což si zaslouží komentář. V předchozích studiích jsme porovnávali oční pohyby u podnětů, které si vzájemně odpovídaly (různé verze slidů v Stárková et al., 2019) nebo dávalo smysl o podobném designu přemýšlet (manipulace směru pohledu na obraze v Kesner et al., 2018). V této studii jsme manipulovali jen kontextovými vodítky (zvýraznění tváře), které účastníci sledovali jen minimálně. Samotné fotografie byly rozmístěny náhodně a tato randomizace je významným rysem použitého designu. Vliv jednotlivých fotografií či míst na obrazovce nás nezajímal a chtěli jsme tyto vlivy randomizací odstínit. Protože nás preference na úrovni jednotlivých obrázků nezajímají, nemělo by smysl oční pohyby spojitě porovnávat a tak používat jemnější způsob hledání rozdílů.

5.4 Vliv expertizy na zrakové vnímání

V předcházejících studiích jsme sledovali rozdíly ve vnímání podnětů, ale účastníky výzkumu jsme považovali za homogenní skupinu. V některých studiích účastníci viděli jen jeden typ podnětů (mezisubjektový design v Stárková et al., 2019), v dalších viděli více variant podnětů a mohli jsme tak statisticky porovnávat výsledky na úrovni jednotlivých osob (Kesner et al., 2018).

Řada studií se zaměřuje na porovnání různých skupin jako jsou zdraví a nemocní nebo staří a mladí lidé. V naší čtvrté studii (Adámek et al., 2019) jsme kombinovali oba přístupy. Každý podnět jsme měli připravený ve dvou variantách: jednou jako fotografií a jednou jako obraz, který malíř podle fotografie vytvořil. Zajímalo nás, do jaké míry se budou oční pohyby u obou verzí podobat. Předpokládali jsme, že se při tvorbě obrazu umělec nesnaží o věrnou kopii, ale vkládá do tohoto procesu své záměry. Tím, že některé aspekty obsahu zvýrazňuje a jiné tlumí, manipuluje s pozorností diváka a provází jej obsahem obrazu. Kromě sledování uměleckého záměru nás zajímalo, zda se bude lišit vnímání obrazů a fotografií u laiků na jedné straně a u zkušených pozorovatelů (výtvarníci, historici umění a pokročilí studenti těchto oborů) na straně druhé.

Našemu týmu se podařilo shromáždit jedinečný soubor 90 fotografií a 90 obrazů se stejným obsahem. Tyto podněty jsme prezentovali 21 laikům a 19 odborníkům. Experiment se skládal ze dvou sezení a každý podnět byl v daném sezení vždy prezentován buď jen jako obraz nebo jako fotografie. Mezi oběma sezeními byl alespoň dvoutýdenní odstup a v druhém sezení byly podněty prezentovány v alternativní variantě (původní obraz jako fotografie a naopak). Každý podnět si účastníci prohlíželi po dobu 10 sekund bez konkrétního zadání.

Oční pohyby jsme porovnávali na dvou úrovních. Na jednoduché úrovni nás zajímali změny v délce sakád a trvání fixací. Na složitější úrovni nás zajímala distribuce pohledu. Podnět jsme si pomyslně pokryli čtvercovou sítí a měřili, kolik polí sítě (procent plochy) pozorovatel navštívil pohledem. Distribuce pohledu je možné porovnávat a měřit tak podobnost mezi obrazy a fotografiemi nebo laiky a experty. Jako konkrétní míru podobnosti jsme používali AUC (area under curve), která postihuje, jak přesně se dá distribuce pohledů jedné skupiny předvídat z distribuce druhé skupiny. Podobný způsob měření byl již používán v jiných předchozích studiích (např. Fuchs et al., 2011).

Zjistili jsme, že při prohlížení obrazů dělají pozorovatelé více kratších fixací, které jsou v průměru dál od sebe a prozkoumají větší plochu obrazu oproti fotografiím. Experti celkově prohlíželi větší plochu u obou druhů podnětů a rozdíl byl patrnější u obrazů. Podobnost distribuce pohledů u fotografií a u obrazů byla poměrně velká. U expertů byla podobnost o trochu menší, což pravděpodobně souvisí s mírně odlišným způsobem prohlížení obrazů a s větším množstvím prozkoumané plochy.

V této studii jsme porovnávali podobnost očních pohybů. Použitý způsob připomíná metody popisované v předchozích studiích o MOT (NSS, korelační vzdálenost). Na rozdíl od těchto metod se jedná o diskrétní kategorizaci, tj. porovnáваме, zda pohledy padly do stejného čtverce použité sítě bez ohledu na jejich vzdálenost. Metody NSS a korelační vzdálenosti používají v jedné fázi podobná data, ale liší se závěrečným použitím gaussovského filtru, který pomůže *spojitě započítat údaje o vzdálenosti* mezi jednotlivými čtverci. Síť používaná pro klasifikaci podnětů bývá také jemnější.

Důležitým rozšířením studie by mohlo být *testování míry šumu v datech*. V našem experimentu jsme sice zaznamenali statisticky významné odlišnosti v distribuci pohledu expertů a laiků, ale bylo by užitečné vědět, o jak velký rozdíl se jedná v kontextu celkové variability, kterou napříč lidmi vidíme. S ohledem na povahu úlohy by nemělo cenu měřit míru šumu pomocí opakovaných prezentací stejným účastníkům (jako jsme mohli dělat u MOT), ale základní představu by mohlo přinést i srovnávání napříč účastníky. Mohli bychom tak zkoumat, zda odlišné chování expertů je konsistentní (přitahují je drobnosti, kterých si laici nevšimnou) a nebo není (chtějí prozkoumat větší část obrazu, ale vzájemně se liší v tom, co upoutá jejich pozornost).

5.5 Shrnutí potenciálu metod

Přestože se předkládané studie liší od dynamických úloh používaných v představených metodologických studiích, některé postupy by se daly s výhodou aplikovat. Častým omezením v analýze dat byla potřeba kategorizace pohledů. Buď binární rozlišení, zda pohled padá do některých oblastí zájmu (AOI), nebo zúžení analýzy na nějaký úvodní časový interval. Jak jsem ukazoval dříve, tato diskretizace přináší problémy, protože lidé nepotřebují směřovat svůj pohled vždy přesně nebo protože pohled může na změnu obsahu reagovat se zpožděním. V řadě úloh proto může pomoci spojité hodnocení podobnosti (NSS, korelační vzdálenost).

Míru podobnosti resp. nepodobnosti můžeme následně statisticky testovat a zohledňovat celkovou variabilitu pohledů (Děchtěrenko et al., 2017). Tato variabilita bývá jinak ignorována a projevuje se jen jako obecný šum ve statistické analýze.

V některých případech nás může zajímat právě podobnost pohledů (rozdíly mezi skupinami, rozdíly mezi druhy podnětů) a tato data spojitým hodnocením přímo dostáváme. Nemusíme se pak spoléhat na nepřímé indikátory podobnosti, kdy prohlížení považujeme za podobné, pokud pozorovatel navštíví určité oblasti v podobném poměru.

Jednotlivé studie pracovaly se statickými podněty, ale výhodou představených metod je snadné rozšíření o změny v čase, což otevírá další možnosti výzkumu.

Kapitola 6

Eye movements in repeated multiple object tracking

Jiří Lukavský

Publikováno v *Journal of Vision*, 13(7), 1–16.

DOI: 10.1167/13.7.9

Datum publikace 17.6.2013

Impact Factor (2013): 2.727

© 2013 ARVO

Kapitola 7

Gaze position lagging behind scene content in multiple object tracking: Evidence from forward and backward presentations

Jiří Lukavský

Filip Děchtěrenko

Publikováno v *Attention, Perception, & Psychophysics*, 78(8), 2456–2468.

DOI: 10.3758/s13414-016-1178-4

Datum publikace 26.7.2016

Impact Factor (2016): 1.863

© 2016 The Psychonomic Society, Inc.

Kapitola 8

Flipping the stimulus: Effects on scanpath coherence?

Filip Děchtěrenko

Jiří Lukavský

Kenneth Holmqvist

Publikováno v *Behavior Research Methods*, 49(1), 382–393.

DOI: 10.3758/s13428-016-0708-2

Datum publikace 2.3.2016

Impact Factor (2016): 3.623

© 2016 The Psychonomic Society, Inc.

Kapitola 9

Scene categorization in the presence of a distractor

Jiří Lukavský

Publikováno v *Journal of Vision*, 19(2), 6–6.

DOI: 10.1167/19.2.6

Datum publikace 8.2.2019

Impact Factor (2019): 2.145

Creative Commons Attribution 4.0 International License

Kapitola 10

Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning?

Tereza Stárková

Jiří Lukavský

Ondřej Javora

Cyril Brom

Publikováno v *Journal of Computer Assisted Learning*, 35(4), 555-568.

DOI: 10.1111/jcal.12359

Datum publikace 7.5.2019

Impact Factor (2019): 2.126

© 2019 John Wiley & Sons Ltd

Kapitola 11

Perception of direct vs. averted gaze in portrait paintings: An fMRI and eye-tracking study

Ladislav Kesner

Dominika Grygarová

Iveta Fajnerová

Jiří Lukavský

Tereza Nekovářová

Jaroslav Tintěra

Yuliya Zaytseva

Jiří Horáček

Publikováno v *Brain and Cognition*, 125, 88-99.

DOI: 10.1016/j.bandc.2018.06.004

Datum publikace 15.6.2018

Impact Factor (2018): 2.619

© 2018 Elsevier Inc.

Kapitola 12

illuminating Smiles and Frowns: Visual-Affective Cueing Influences Viewer Perceptions of Page Layout Images

Dominika Grygarová

Petr Adámek

Jiří Lukavský

Ladislav Kesner

Publikováno v *Perceptual and Motor Skills*, 127(1), 75-97.

DOI: 10.1177/0031512519876742

Datum publikace 23.9.2019

Impact Factor (2019): 1.245

© 2019 The Author(s)

Kapitola 13

Tracking artistic transformations: comparing paintings and their source photographs

Petr Adámek

Dominika Grygarová

Jiří Lukavský

Ladislav Kesner

Publikováno v *Československá psychologie*, 63(3), 337-353.

Datum publikace 2019

Impact Factor (2019): 0.478

© 2019 Československá psychologie

Kapitola 14

Podíl autora na přiložených studiích

Provedené výzkumy a prezentované práce by nebyly možné bez přispění mých kolegů, kterým bych chtěl na tomto místě poděkovat.

- 1) Lukavský, J. (2013). **Eye movements in repeated multiple object tracking.** *Journal of Vision*, 13(7), 1–16.

Byl jsem jediným autorem této experimentální studie.

Autorský podíl dle CRediT¹: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing

- 2) Lukavský, J., & Děchtěrenko, F. (2016). **Gaze position lagging behind scene content in multiple object tracking: Evidence from forward and backward presentations.** *Attention, Perception, & Psychophysics*, 78(8), 2456–2468.

Podílel jsem se na všech výzkumných fázích projektu.

Podíl: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing

- 3) Děchtěrenko, F., Lukavský, J., & Holmqvist, K. (2017). **Flipping the stimulus: Effects on scanpath coherence?** *Behavior Research Methods*, 49(1), 382–393.

Podílel jsem se hlavně na konceptuálním plánování studie a psaní textu. Byl jsem vedoucím dizertační práce FD.

Podíl: Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing

¹Pro popis podílu jsou použity CRediT, což je taxonomie autorských rolí používaná některými vydavateli. Více na <https://casrai.org/credit/>

- 4) Lukavský, J. (2019). Scene categorization in the presence of a distractor. *Journal of Vision*, 19(2), 6–6.

Byl jsem jediným autorem této experimentální studie.

Podíl: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing

- 5) Stárková, T., Lukavský, J., Javora, O., & Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning? *Journal of Computer Assisted Learning*, 35(4), 555–568.

Podílel jsem se na technické a metodologické přípravě měření očních pohybů, jejich analýze a psaní textu.

Podíl: Formal Analysis, Methodology, Software, Writing – original draft, Writing – review & editing

- 6) Kesner, L., Grygarová, D., Fajnerová, I., Lukavský, J., Nekovářová, T., Tintěra, J., Zaytseva, Y., & Horáček, J. (2018). Perception of direct vs. averted gaze in portrait paintings: An fMRI and eye-tracking study. *Brain and Cognition*, 125, 88–99.

Podílel jsem se na technické a metodologické přípravě měření očních pohybů, jejich analýze a psaní textu.

Podíl: Formal Analysis, Methodology, Software, Writing – original draft, Writing – review & editing

- 7) Grygarová, D., Adámek, P., Lukavský, J., & Kesner, L. (2020). Illuminating Smiles and Frowns: Visual-Affective Cueing Influences Viewer Perceptions of Page Layout Images. *Perceptual and Motor Skills*, 127(1), 75–97.

Podílel jsem se na metodologické přípravě měření očních pohybů, jejich analýze a psaní textu.

Podíl: Formal Analysis, Methodology, Writing – original draft, Writing – review & editing

- 8) Adámek, P., Grygarová, D., Lukavský, J., & Kesner, L. (2019). Tracking artistic transformations: comparing paintings and their source photographs. *Československá Psychologie*, 63(3), 337–353.

Podílel jsem se na metodologické přípravě měření očních pohybů, jejich analýze a psaní textu.

Podíl: Formal Analysis, Methodology, Writing – original draft, Writing – review & editing

Kapitola 15

Literatura

- Adámek, P., Grygarová, D., Lukavský, J., & Kesner, L. (2019). Tracking artistic transformations: Comparing paintings and their source photographs. *Ceskoslovenska Psychologie*, *63*(3), 337–353.
- Allen, R., McGeorge, P., Pearson, D., & Milne, A. B. (2004). Attention and expertise in multiple target tracking. *Applied Cognitive Psychology*, *18*(3), 337–347. <https://doi.org/10.1002/acp.975>
- Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science*, *16*(8), 637–643. <https://doi.org/10.1111/j.1467-9280.2005.01587.x>
- Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(13). <https://doi.org/10.1167/7.13.14>
- Alvarez, G. A., & Oliva, A. (2008). The Representation of Simple Ensemble Visual Features Outside the Focus of Attention. *Psychological Science*, *19*(4), 392–398.
- Anderson, B. (2011). There is no Such Thing as Attention. *Frontiers in Psychology*, *2*. <https://doi.org/10.3389/fpsyg.2011.00246>
- Atsma, J., Koning, A., & van Lier, R. (2012). Multiple object tracking: Anticipatory attention doesn't "bounce". *Journal of Vision*, *12*(1). <https://doi.org/10.1167/12.13.1>
- Brom, C., Stárková, T., Lukavský, J., Javora, O., & Bromová, E. (2016). Eye tracking in emotional design research: What are its limitations? *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, 114:1–114:6.
- Bylinskii, Z., Recasens, A., Borji, A., Oliva, A., Torralba, A., & Durand, F. (2016). Where Should Saliency Models Look Next? *Computer Vision ECCV 2016*, 809–824. https://doi.org/10.1007/978-3-319-46454-1_49
- Carlson, T. A., Alvarez, G. A., & Cavanagh, P. (2007). Quadrantic deficit reveals anatomical constraints on selection. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(33), 13496–13500. <https://doi.org/http://dx.doi.org/10.1073/pnas.0702685104>

- Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, 9(7), 349–354.
- Chun, M. M., & Jiang, Y. (1998). Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. *Cognitive Psychology*, 36(1), 28–71. <https://doi.org/10.1006/cogp.1998.0681>
- Davenport, J. L. (2007). Consistency effects between objects in scenes. *Memory & Cognition*, 35(3), 393–401. <https://doi.org/10.3758/BF03193280>
- Davenport, J. L., & Potter, M. C. (2004). Scene Consistency in Object and Background Perception. *Psychological Science*, 15(8), 559–564. <https://doi.org/10.1111/j.0956-7976.2004.00719.x>
- DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, 17(6-7), 790–811. <https://doi.org/10.1080/13506280902793843>
- Dechterenko, F., & Lukavský, J. (2018). Robustness of Metrics Used for Scanpath Comparison. *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, 63:1–63:5. <https://doi.org/10.1145/3204493.3204580>
- Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior Research Methods*, 44(4), 1079–1100. <https://doi.org/10.3758/s13428-012-0212-2>
- Děchtěrenko, F., & Lukavský, J. (2016). Predicting Eye Movements in Multiple Object Tracking Using Neural Networks. *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, 271–274. <https://doi.org/10.1145/2857491.2857502>
- Děchtěrenko, F., Lukavský, J., & Holmqvist, K. (2017). Flipping the stimulus: Effects on scanpath coherence? *Behavior Research Methods*, 49(1), 382–393. <https://doi.org/10.3758/s13428-016-0708-2>
- Dickinson, C., & Intraub, H. (2009). Spatial Asymmetries in Viewing and Remembering Scenes: Consequences of an Attentional Bias? *Attention, Perception & Psychophysics*, 71(6), 1251–1262.
- Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision*, 10(10), 28. <https://doi.org/10.1167/10.10.28>
- Eysenck, M. W., & Keane, M. T. (2008). *Kognitivní psychologie*. Academia.
- Fehd, H. M., & Seiffert, A. E. (2008). Eye movements during multiple object tracking: Where do participants look? *Cognition*, 108(1), 201–209. <https://doi.org/10.1016/j.cognition.2007.11.008>
- Fehd, H. M., & Seiffert, A. E. (2010). Looking at the center of the targets helps multiple object tracking. *Journal of Vision*, 10(19). <https://doi.org/10.1167/10.4.19>
- Fencsik, D. E., Klieger, S. B., & Horowitz, T. S. (2007). The role of location and motion information in the tracking and recovery of moving objects. *Attention, Perception, & Psychophysics*, 69(4), 567–577. <https://doi.org/10.3758/BF03193914>

- Feusner, M., & Lukoff, B. (2008). Testing for statistically significant differences between groups of scan patterns. *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications - ETRA '08*, 43. <https://doi.org/10.1145/1344471.1344481>
- Foulsham, T., & Kingstone, A. (2013). Optimal and Preferred Eye Landing Positions in Objects and Scenes. *Quarterly Journal of Experimental Psychology*, 66(9), 1707–1728. <https://doi.org/10.1080/17470218.2012.762798>
- Franconeri, S. L., Alvarez, G. A., & Enns, J. T. (2007). How many locations can be selected at once? *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1003–1012. <https://doi.org/10.1037/0096-1523.33.5.1003>
- Franconeri, S. L., Jonathan, S. V., & Scimeca, J. M. (2010). Tracking Multiple Objects Is Limited Only by Object Spacing, Not by Speed, Time, or Capacity. *Psychological Science*, 21(7), 920–925. <https://doi.org/10.1177/0956797610373935>
- Franconeri, S. L., Pylyshyn, Z. W., & Scholl, B. J. (2012). A simple proximity heuristic allows tracking of multiple objects through occlusion. *Attention, Perception, & Psychophysics*, 74(4), 691–702. <https://doi.org/10.3758/s13414-011-0265-9>
- Fuchs, I., Ansorge, U., Redies, C., & Leder, H. (2011). Saliency in paintings: Bottom-Up influences on eye fixations. *Cognit. Comput.*, 3(1), 25–36.
- Greene, M. R., & Fei-Fei, L. (2014). Visual categorization is automatic and obligatory: Evidence from Stroop-like paradigm. *Journal of Vision*, 14(14). <https://doi.org/10.1167/14.1.14>
- Grygarová, D., Adámek, P., Lukavský, J., & Kesner, L. (2019). Illuminating smiles and frowns: Visual-Affective cueing influences viewer perceptions of page layout images. *Percept. Mot. Skills*, 0031512519876742.
- Hagenbeek, R. E., & Van Strien, J. W. (2002). LeftRight and UpperLower Visual Field Asymmetries for Face Matching, Letter Naming, and Lexical Decision. *Brain and Cognition*, 49(1), 34–44. <https://doi.org/10.1006/brcg.2001.1481>
- Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour*, 1. <https://doi.org/10.1038/s41562-017-0208-0>
- Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J.-H., & Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-019-01846-w>
- Hopf, J.-M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H.-J., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences*, 103(4), 1053–1058. <https://doi.org/10.1073/pnas.0507746103>
- Horowitz, T. S., & Cohen, M. A. (2010). Direction information in multiple object tracking is limited by a graded resource. *Attention, Perception & Psychophysics*, 72(7), 1765–1775. <https://doi.org/10.3758/APP.72.7.1765>

- Howard, C. J., & Holcombe, A. O. (2008). Tracking the changing features of multiple objects: Progressively poorer perceptual precision and progressively greater perceptual lag. *Vision Research*, *48*(9), 1164–1180. <https://doi.org/10.1016/j.visres.2008.01.023>
- Howard, C. J., Masom, D., & Holcombe, A. O. (2011). Position representations lag behind targets in multiple object tracking. *Vision Research*, *51*(17), 1907–1919. <https://doi.org/10.1016/j.visres.2011.07.001>
- Howard, C. J., Uttley, J., & Andrews, S. (2018). Chapter 4 - Team ball sport participation is associated with performance in two sustained visual attention tasks: Position monitoring and target identification in rapid serial visual presentation streams. In *Progress in Brain Research* (Vol. 240, pp. 53–69). Elsevier. <https://doi.org/10.1016/bs.pbr.2018.09.001>
- Huff, M., Papenmeier, F., Jahn, G., & Hesse, F. W. (2010). Eye movements across viewpoint changes in multiple object tracking. *Visual Cognition*, *18*(9), 1368–1391. <https://doi.org/10.1080/13506285.2010.495878>
- Hyönä, J., Li, J., & Oksama, L. (2019). Eye Behavior During Multiple Object Tracking and Multiple Identity Tracking. *Vision*, *3*(3), 37. <https://doi.org/10.3390/vision3030037>
- Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of attention and monitoring of velocities during multiple-object tracking. *Journal of Vision*, *9*(1). <https://doi.org/10.1167/9.4.1>
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259. <https://doi.org/10.1109/34.730558>
- James, W. (1890). *The principles of psychology*. Henry Holt.
- Jarodzka, H., Holmqvist, K., & Nyström, M. (2010). A vector-based, multidimensional scanpath similarity measure. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10*, 211. <https://doi.org/10.1145/1743666.1743718>
- Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*(4), 346–368. <https://doi.org/10.1016/j.cogpsych.2005.12.001>
- Kennedy, D. P., D'Onofrio, B. M., Quinn, P. D., Bölte, S., Lichtenstein, P., & Falck-Ytter, T. (2017). Genetic Influence on Eye Movements to Complex Scenes at Short Timescales. *Current Biology*, *27*(22), 3554–3560.e3. <https://doi.org/10.1016/j.cub.2017.10.007>
- Kesner, L., Grygarová, D., Fajnerová, I., Lukavský, J., Nekovářová, T., Tintěra, J., Zaytseva, Y., & Horáček, J. (2018). Perception of direct vs. Averted gaze in portrait paintings: An fMRI and eye-tracking study. *Brain Cogn.*, *125*, 88–99.
- Krauzlis, R. J., Bollimunta, A., Arcizet, F., & Wang, L. (2014). Attention as an effect not a cause. *Trends in Cognitive Sciences*, *18*(9), 457–464.

- Kummerer, M., Wallis, T. S. A., Gatys, L. A., & Bethge, M. (2017). Understanding Low- and High-Level Contributions to Fixation Prediction. *Proceedings of the IEEE International Conference on Computer Vision*, 4789–4798.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* (Technical Report A-8). University of Florida.
- Larson, A. M., Freeman, T. E., Ringer, R. V., & Loschky, L. C. (2014). The spatiotemporal dynamics of scene gist recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 40(2), 471–487. <https://doi.org/10.1037/a0034986>
- Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, 9(6). <https://doi.org/10.1167/9.10.6>
- Loschky, L. C., Hansen, B. C., Sethi, A., & Pydimarri, T. N. (2010). The role of higher order image statistics in masking scene gist recognition. *Attention, Perception, & Psychophysics*, 72(2), 427–444. <https://doi.org/10.3758/APP.72.2.427>
- Lukavský, J. (2019). Scene categorization in the presence of a distractor. *Journal of Vision*, 19(2), 6–6. <https://doi.org/10.1167/19.2.6>
- Lukavský, J. (2013). Eye movements in repeated multiple object tracking. *Journal of Vision*, 13(7), 1–16. <https://doi.org/10.1167/13.7.9>
- Lukavský, J., & Děchtěrenko, F. (2016). Gaze position lagging behind scene content in multiple object tracking: Evidence from forward and backward presentations. *Attention, Perception, & Psychophysics*, 78(8), 2456–2468. <https://doi.org/10.3758/s13414-016-1178-4>
- Luke, S. G., Darowski, E. S., & Gale, S. D. (2018). Predicting eye-movement characteristics across multiple tasks from working memory and executive control. *Memory & Cognition*. <https://doi.org/10.3758/s13421-018-0798-4>
- Mayer, R. E. (2009). *Multimedia learning*. Cambridge University Press. <https://books.google.cz/books?id=PSSus6qEHpMC>
- Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, 33, 12–18.
- Merkel, C., Hopf, J.-M., & Schoenfeld, M. A. (2017). Spatio-temporal dynamics of attentional selection stages during multiple object tracking. *NeuroImage*, 146, 484–491. <https://doi.org/10.1016/j.neuroimage.2016.10.046>
- Meyerhoff, H. S., Papenmeier, F., & Huff, M. (2017). Studying visual attention using the multiple object tracking paradigm: A tutorial review. *Attention, Perception, & Psychophysics*, 1–20. <https://doi.org/10.3758/s13414-017-1338-1>
- Meyerhoff, H. S., Papenmeier, F., Jahn, G., & Huff, M. (2015). Distractor Locations Influence Multiple Object Tracking Beyond Interobject Spacing: Evidence From Equidistant Distractor Displacements. *Experimental*

- Psychology*, 62, 170–180. <https://doi.org/10.1027/1618-3169/a000283>
- Meyerhoff, H. S., Papenmeier, F., Jahn, G., & Huff, M. (2016). Not FLEXible Enough: Exploring the Temporal Dynamics of Attentional Reallocations With the Multiple Object Tracking Paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 42(6), 776–787. <https://doi.org/10.1037/xhp0000187>
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9(3), 353–383. [https://doi.org/10.1016/0010-0285\(77\)90012-3](https://doi.org/10.1016/0010-0285(77)90012-3)
- Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, 370(6487), 256–257. <https://doi.org/10.1038/370256b0>
- Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences*, 31(2), 179–239. <https://doi.org/10.1017/S0140525X08003804>
- Nijhawan, R., & Khurana, B. (2000). Conscious registration of continuous and discrete visual events. In *Neural correlates of consciousness: Empirical and conceptual questions* (pp. 203–219). The MIT Press.
- Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision*, 10(8), 20–20. <https://doi.org/10.1167/10.8.20>
- Ogawa, H., Watanabe, K., & Yagi, A. (2009). Contextual cueing in multiple object tracking. *Visual Cognition*, 17(8), 1244–1258. <https://doi.org/10.1080/13506280802457176>
- Oliva, A., & Schyns, P. G. (1997). Coarse Blobs or Fine Edges? Evidence That Information Diagnosticity Changes the Perception of Complex Visual Stimuli. *Cognitive Psychology*, 34(1), 72–107. <https://doi.org/10.1006/cogp.1997.0667>
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, 41(2), 176–210. <https://doi.org/10.1006/cogp.1999.0728>
- Peters, R. J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Res.*, 45(18), 2397–2416.
- Portilla, J., & Simoncelli, E. P. (2000). A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients. *International Journal of Computer Vision*, 40(1), 49–70. <https://doi.org/10.1023/A:1026553619983>
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25. <https://doi.org/10.1080/00335558008248231>
- Pylyshyn, Z. W. (1994). Some primitive mechanisms of spatial attention. *Cognition*, 50(1), 363–384. [https://doi.org/10.1016/0010-0277\(94\)90036-1](https://doi.org/10.1016/0010-0277(94)90036-1)
- Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, 32(1), 65–97. [https://doi.org/10.1016/0010-0277\(89\)90014-0](https://doi.org/10.1016/0010-0277(89)90014-0)
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80(1-2), 127–158. [https://doi.org/http://dx.doi.org/10.1016/S0010-0277\(00\)00156-6](https://doi.org/http://dx.doi.org/10.1016/S0010-0277(00)00156-6)

- Pylyshyn, Z. W. (2007). *Things and places: How the mind connects with the world*. MIT Press.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 179–197. <https://doi.org/10.1163/156856888X00122>
- Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 49(2), 151–173. <https://doi.org/10.1037/1196-1961.49.2.151>
- Reingold, E. M., & Sheridan, H. (2011). Eye movements and visual expertise in chess and medicine. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 528–550). Oxford University Press.
- Risko, E. F., Anderson, N. C., Lanthier, S., & Kingstone, A. (2012). Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. *Cognition*, 122(1), 86–90. <https://doi.org/10.1016/j.cognition.2011.08.014>
- Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic Semantic Processing in a Picture-Word Interference Task. *Child Development*, 46(1), 247–253. <https://doi.org/10.2307/1128859>
- Salas, J. A., & Levin, D. T. (2021). Efficient calculations of NSS-based gaze similarity for time-dependent stimuli. *Behav. Res. Methods*.
- Scholl, B. J. (2009). What Have We Learned about Attention from Multiple-Object Tracking (and Vice Versa)? In D. Dedrick & L. Trick (Eds.), *Computation, Cognition, and Pylyshyn* (pp. 49–77). The MIT Press.
- Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking Multiple Items Through Occlusion: Clues to Visual Objecthood. *Cognitive Psychology*, 38(2), 259–290. <https://doi.org/http://dx.doi.org/10.1006/cogp.1998.0698>
- Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, 80(1-2), 159–177.
- Schupp, H. T., Schmälzle, R., Fleisch, T., Weike, A. I., & Hamm, A. O. (2012). Affective picture processing as a function of preceding picture valence: An ERP analysis. *Biological Psychology*, 91(1), 81–87. <https://doi.org/https://doi.org/10.1016/j.biopsycho.2012.04.006>
- Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision*, 11(5), 9–9. <https://doi.org/10.1167/11.5.9>
- Schyns, P. G., & Oliva, A. (1994). From Blobs to Boundary Edges: Evidence for Time- and Spatial-Scale-Dependent Scene Recognition. *Psychological Science*, 5(4), 195–200. <https://doi.org/10.1111/j.1467-9280.1994.tb00500.x>
- Sekuler, R., McLaughlin, C., & Yotsumoto, Y. (2008). Age-Related Changes in Attentional Tracking of Multiple Moving Objects. *Perception*, 37(6), 867–876. <https://doi.org/10.1068/p5923>
- Shim, W. M., Alvarez, G. A., & Jiang, Y. V. (2008). Spatial separation between targets constrains maintenance of attention on multiple objects. *Psychonomic Bulletin & Review*, 15(2), 390–397. <https://doi.org/10.3758/PBR.15.2.390>

- Starreveld, P. A., & Heij, W. L. (2017). Picture-word interference is a Stroop effect: A theoretical analysis and new empirical findings. *Psychonomic Bulletin & Review*, 24(3), 721–733. <https://doi.org/10.3758/s13423-016-1167-6>
- Stárková, T., Lukavský, J., Javora, O., & Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning? *Journal of Computer Assisted Learning*, 35(4), 555–568.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, 18(6), 643–662. <https://doi.org/10.1037/h0054651>
- Suchow, J. W., Fougny, D., Brady, T. F., & Alvarez, G. A. (2014). Terms of the debate on the format and structure of visual memory. *Attention, Perception & Psychophysics*, 76(7), 2071–2079. <https://doi.org/10.3758/s13414-014-0690-7>
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381(6582), 520–522. <https://doi.org/10.1038/381520a0>
- Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems*, 14, 391–412.
- Upadhyayula, A., & Flombaum, J. (2020). A model that adopts human fixations explains individual differences in multiple object tracking. *Cognition*, 104418. <https://doi.org/10.1016/j.cognition.2020.104418>
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, 24(3), 295–340.
- Yarbus, A. L. (1967). *Eye movements and vision*. Springer.
- Zelinsky, G. J., & Neider, M. B. (2008). An eye movement analysis of multiple object tracking in a realistic environment. *Visual Cognition*, 16(5), 553–566. <https://doi.org/10.1080/13506280802000752>
- Zelinsky, G. J., & Todor, A. (2010). The role of “rescue saccades” in tracking objects through occlusions. *Journal of Vision*, 10(14), 29. <https://doi.org/10.1167/10.14.29>