

a odvozovat z nich závěry. Tato publikace by měla posloužit k tomu, aby cesta dalších výzkumníků za výsledky byla pohodlnější, rychlejší a méně trnitá.

První kapitola popisuje, jak technologie eye-tracking funguje. Ve druhé kapitole je stručně shrnuto využití eye-trackingu v různých oblastech lidské činnosti. Třetí kapitola popisuje kognitivní kartografii. Čtvrtá kapitola je zaměřena více prakticky. Jsou v ní popsány tři eye-tracking systémy – SMI RED 250, Gazepoint a EyeTribe v kombinaci s aplikací OGAMA. Po přečtení této kapitoly by měl být čtenář schopen toto zařízení zapojit a nastavit tak, aby mohl začít nahrávat data. Před vlastním záznamem pohybu očí je však důležité rozmyslet si design experimentu a to, jak budou stimuly respondentům prezentovány. Tomu se věnuje pátá kapitola. Šestá kapitola je pak zaměřena na pre-processing dat, který obsahuje identifikaci fixací a sakád a také čištění dat a hledání chyb v datech. Poslední, sedmá kapitola popisuje několik metod analýzy a vizualizace eye-tracking dat, a to včetně statistické analýzy a doporučení pro výběr vhodné metody.

Přestože se v názvu knihy objevuje kognitivní kartografie, lze většinu rad a doporučení v této knize obsažených využít i v jiných oblastech výzkumu.

Standa Popelka

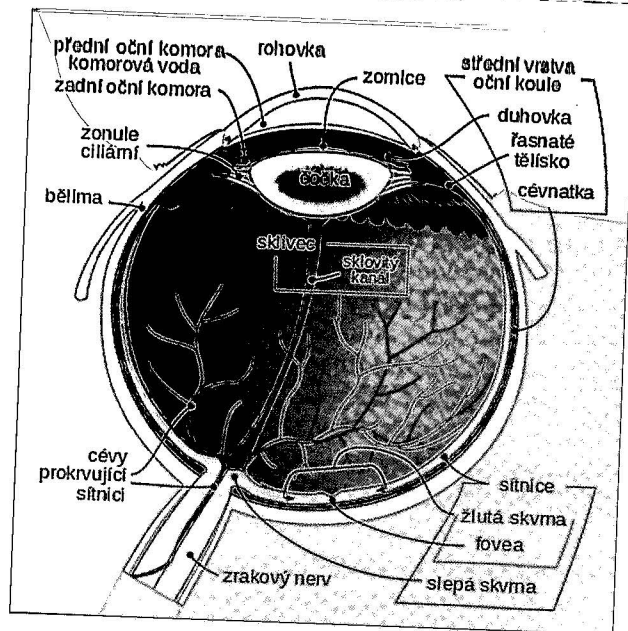
## 1 EYE-TRACKING

Pro sledování pohybu očí neboli eye-tracking je nutné nejprve znát několik informací o tom, jak vlastně lidský zrak funguje a jaké pohyby očí vykonávají.

### 1.1 Charakteristika lidského oka

Zrak je u lidí nejsilnější a nejvíce využívaný smysl. Obě oči poskytují zhruba eliptický obraz prostoru o úhlovém rozpětí přibližně  $200^\circ$  na šířku a  $130^\circ$  na výšku (Biedert a kol., 2010). Struktura lidského oka je uzpůsobena potřebě promítnout paprsek světla na sítnici. Oko propouští světlo skrz zornici, následně je obraz převrácen a zaostřen na zadní stranu oční bulvy – sítnici. Světlo dopadající na sítnici pak způsobuje chemické přeměny ve světločivých buňkách. Těmi jsou tyčinky a čípky, které přeměnění dopadající světlo v elektrické impulzy, které jsou pomocí očního nervu přeneseny do mozku.

Oční bulbus má přibližně kulovitý tvar o průměru asi 23 mm. V oční štěrbině mezi otevřenými víčky je viditelný pouze menší, přední úsek oční koule, zatímco větší, zadní část je uložena v hloubi oční. Oční koule obsahuje přední pól, který odpovídá vrcholu rohovky, a zadní pól. Oba póly spojuje oční (geometrická) osa. Tato geometrická osa není totožná se zornou (vizuální) osou oka spojující optický střed oka se žlutou skvrnou. Geometrická a vizuální osa spolu svírají úhel asi  $5^\circ$ . Oční koule je tvořena stěnou oční koule a obsahem oční koule. Stěna oční koule se skládá ze zevní vazivové vrstvy (bělma a rohovka), střední vrstvy (cévnatka, řasnaté tělísko, duhovka) a vnitřní vrstvy (sítnice). Obsah oční koule představují průhledné a čiré struktury, které propouštějí světelné paprsky a soustřeďují je (lámou) tak, že dopadají na sítnici. Jsou to optická prostředí oka a tvoří je čočka, sklívec a přední a zadní oční komora (Synek a Skorovská, 2014). Struktura lidského oka je znázorněna na obrázku 1. Na následujících stranách budou popsány jednotlivé části oka a jejich význam při eye-trackingu.



Obr. 1 Struktura lidského oka

**Řasnaté těleso** (corpus ciliare) má tvar zřaseného prstence, který je přiložen na vnitřní stranu bělimy. Na řezu má tvar trojúhelníku. Zadní okraj řasnatého tělesa se ztenčuje a plynule přechází do cévnatky. Směrem dopředu se řasnaté těleso ztlušťuje a spojuje se s duhovkou. Podkladem řasnatého tělesa je vazivové stroma, v němž jsou uloženy četné hladké svalové buňky tvořící musculus ciliaris. Musculus ciliaris tvoří prstenčitý svěrač, který svou kontrakcí uvolňuje závěsný aparát čočky. Ta svou pružností mění tvar a optickou mohutnost, a tím umožňuje akomodaci čočky (Synek a Skorovská, 2014).

**Duhovka** (iris) tvoří nejvíce dopředu vysunutou část prostřední vrstvy oční stěny. Má tvar mezikruží s centrálně uloženým otvorem zvaným zornice neboli panenka (pupilla). Přední plocha duhovky je podle množství pigmentu individuálně různě zbarvená a podmiňuje barvu očí. Duhovka plní v oku úlohu světelné clony. Podle momentálního množství světla si sítnice prostřednictvím pupilárního reflexu mění své osvětlení tak, aby dosahovalo optimálních hodnot (Synek a Skorovská, 2014).

**Sítnice** (retina) tvoří vnitřní vrstvu oka a je dělena na dva stavebně i funkčně odlišné oddíly – optickou a slepou část sítnice. Optická část sítnice je funkčně nejvýznamnější vrstvou stěny oční koule. Na očním pozadí se rozlišuje několik útvarů. U zadního pólu oka leží žlutá skvrna kruhovitě až eliptického tvaru o průměru asi 3 mm. V jejím středu je prohlubeň, která se nazývá fovea centralis. Toto místo leží ve vrcholu optické osy oka a představuje místo nejostřejšího vidění, protože se do něj promítá centrální paprsek. V oblasti žluté skvrny jsou nakupeny jen čípky, které zde mají největší hustotu. Ve vzdálenosti asi čtyř milimetrů od žluté skvrny se vyklenuje bělavé políčko, kterým odstupuje z oční koule zrakový nerv. Toto políčko má tvar kruhu o průměru asi 1,5 mm a neobsahuje žádné světločivé elementy. Představuje tedy fyziologickou slepou skvrnu.

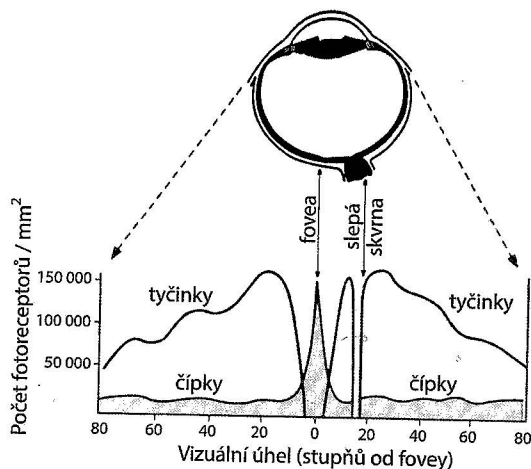
Zevní vrstva sítnice spolu s pigmentovou vrstvou cévnatky působí jako světelná izolační vrstva, která pohlcuje dopadající světelné paprsky a zabráňuje jejich odrazu uvnitř oka („černá komora“). Vnitřní vrstva obsahuje vlastní primární smyslové buňky (fotoreceptory) a neurony, které sbírají informace z fotoreceptorů a odvádějí je prostřednictvím zrakového nervu z oka do mozku. Smyslové buňky se v sítnici vyskytují ve dvojí podobě: tyčinky (bacilli) a čípky (coni). Tyčinky registrují množství dopadajícího světla a představují většinu smyslových buněk sítnice – jejich celkový počet je asi 130 milionů. Čípky slouží k vnímání barev a je jich jen okolo sedmi milionů (Synek a Skorovská, 2014). Čípky jsou hojně zastoupeny především v oblasti fovey, která nepokrývá ani 2 % zorného pole. Z toho důvodu lidské oko vidí ostře pouze velmi malou část zorného pole. Hustota fotoreceptorů (tyčinek a čípků) na sítnici je znázorněna na obrázku 2.

**Bělima** (sclera) je pevná, tuhá vazivová blána, která zaujímá zadních pět šestin oční koule. Je ochranným obalem pro hlubší oddíly oční koule a představuje pevnou oporu pro připojující se šlachy okohybných svalů. Na přední straně bělimy je velký otvor (o průměru asi 12 mm), do jehož okraje (limbus sclerae) je vsazen lehce zkosený okraj rohovky (limbus corneae) (Synek a Skorovská, 2014).

**Rohovka** (cornea) překlenuje a uzavírá kruhový otvor o průměru asi 12 mm v předním okraji bělimy. Má tvar segmentu koule, který se vyklenuje dopředu. Rohovka představuje vstupní oddíl optického prostředí oka a z hlediska indexu lomu světla je jeho nejvýznamnější součástí (Synek a Skorovská, 2014).

**Cévnatka** (choroidea) představuje nejrozsáhlejší část prostřední vrstvy stěny oční koule, protože zaujímá její zadní dvě třetiny. Má podobu tenké černohnědé blány bohaté na cévy, která se povrchovou stranou přikládá k bělimě. Dutinová strana cévnatky je hladká a naléhá na ni sítnice. Cévnatka pak představuje tmavou komoru pro optickou vrstvu sítnice a navíc jsou kapilární sítě vyživovány pigmentové buňky sítnice, tedy čípky a tyčinky. Zadu v cévnatce je otvor, kterým prostupují vlákna zrakového nervu, sítníková tepna a žíla (Synek a Skorovská, 2014).

<https://commons.wikimedia.org/w/index.php?curid=7219151>



Obr. 2 Hustota tyčinek a čípků na sítnici lidského oka (upraveno dle Snowden a kol., 2012)

Fovea je ohraničena parafoveálním pásem a vnější perifoveální oblastí (Iwasaki a Inomata, 1986). Vnímání obrazu těmito částmi sítnice se nazývá periferní vidění, které je v porovnání s foveálním viděním rozmazané.

Obsah oční koule tvoří čočka, sklivec a obsah přední a zadní oční komory vyplněné komorovým mokem.

**Čočka** (lens) má tvar bikonvexní spojné čočky o průměru mezi 9 a 10 mm. Na čočku jsou připojena vlákna závěsného aparátu, kterými se na čočku přenášejí účinky ciliárního svalu. Při pohledu do dálky jsou vlákna řasnatého tělíska napjatá, díky čemuž je čočka zploštělá a má nízkou optickou mohutnost. Při pohledu na blízko se stáhnou ciliární svaly, ciliární aparát se poněkud posune a ochabí svůj tah na čočku. Tím se čočka vyklene (akomoduje) a zvýší svou optickou mohutnost (Synek a Skorovská, 2014).

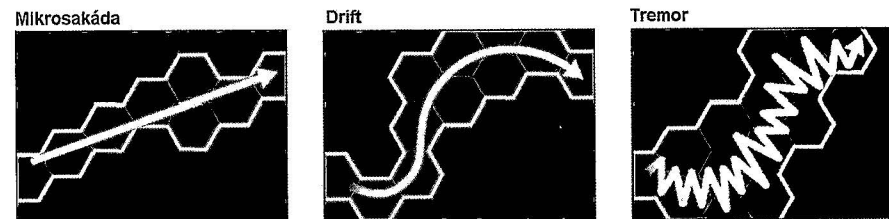
**Sklivec** (corpus vitreum) je měkká huspeninová a průhledná hmota vyplňující sklivcovou komoru. Na přední ploše sklivce je prohloubená jamka, do které je vsazena zadní plocha čočky. Během vývoje probíhá sklivcem k vyvíjející se čočce větev tepny, která však ještě před narozením mizí a ve sklivci po ní zůstává sklovitý kanál nebo jeho úseky (Synek a Skorovská, 2014).

**Oční komory** (camerae bulbi) jsou dva štěrbinové otvory mezi zadní plochou rohovky a přední plochou čočky, jejím závěsným aparátem a řasnatým tělesem. Obě komory jsou vyplněné komorovým mokem, který ze zadní komory proniká přes zornici do přední, odkud je dále odváděn do tzv. Schlemmova kanálu. Při poruše odtoku komorového moku do žilního systému se zvyšuje nitrooční tlak a vzniká tzv. zelený zákal (Synek a Skorovská, 2014).

Pro sledování pohybu očí je klíčové, že více než 40 % lidského mozku zpracovává informace z oblasti pokrývající méně než 5 % zorného pole (Hoffman, 2000; Ware, 2010). Z toho důvodu se oči neustále pohybují tak, aby předměty zájmu byly zobrazeny v oblasti největšího vizuálního rozlišení, tedy fovey. Oční pohyby jsou detailně popsány v následující kapitole.

## 1.2 Pohyby očí – fixace a sakády

Jeden z nejdůležitějších pohybů očí není ve své podstatě pohyb. Jedná se spíše o schopnost udržet oko zaměřené na určitý bod. Tento pohyb oka se označuje jako fixace a dle Holmqvista a kol. (2011) trvá od několika desítek milisekund až po několik sekund. Vizuální vjem se skládá z řady fixací zaměřených na jednotlivé prvky sledované scény. Označení „fixace“ však není úplně přesné, protože oko nikdy není úplně pevně fixováno. Během fixací provádí tři druhy takzvaných mikropohybů. Těmi jsou tremor, mikrosakády a drift (obr. 3).

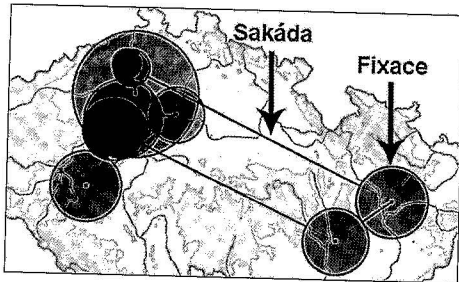


Obr. 3 Mikropohyby probíhající během fixací (Martinez-Conde a Macknik, 2008)

Tremor neboli třes je drobný pohyb s frekvencí okolo 90 Hz, jehož přesný funkční význam není dosud znám. Je však možné, že je způsoben nepřesnou kontrolou okohybných svalů. Jako drift jsou označovány pomalé klouzavé pohyby, které posouvají oko mimo střed fixace. Za 200 ms se osa pohledu může vychýlit až o 6 úhlových minut, a tím pádem se obraz na sítnici posune v rozsahu 10–15 čípků (Synek a Skorovská, 2014). Role mikrosakád spočívá ve vrácení oka zpět do středu fixace. Funkcí těchto pohybů je zabránit efektu saturace receptorů na sítnici, který by vedl ke slábnutí vidění. Lidé si těchto drobných pohybů nejsou vědomi (Biedert a kol., 2010). Mikropohyby probíhající během fixací je možné zaznamenat pouze s využitím vysokofrekvenčních eye-trackerů a jejich studiem se zabývá neurologie. Lze je využít například v klinických studiích zabývajících se autismem (van der Geest a kol., 2001). Pro kartografický výzkum však jejich studium nemá žádný význam, proto jsou analyzovány pouze fixace jako celek.

Pro přesun z jedné fixace na další provádějí oči rychlé balistické pohyby označované jako **sakády**. Sakády představují nejrychlejší pohyb částí lidského těla. Rotační rychlost velkých sakád dosahuje až 500°/s a dle

Holmqvista a kol. (2011) je běžná délka sakád 30–80 ms. Během sakád lidský mozek nevnímá téměř žádné vizuální vjemy. Tento fakt není způsoben pouze rozmazáním vnímaného obrazu, ale také nervovým procesem označovaným jako sakadické potlačení neboli saccadic suppression (Hammoud a Mulligan, 2008). Díky tomu například není možné vidět pohyb vlastních očí v zrcadle. Sakády jsou běžně znázorňovány jako přímé spojnice mezi fixacemi (viz obr. 4). Jen zřídka je však průběh sakád opravdu přímý. Ve skutečnosti mohou sakády nabývat mnoha různých tvarů a zakřivení. Velká část sakád navíc neskončí přesně v cíli (ve středu následující fixace), ale pohyb oka se před zastavením rozkolísá. Tento postsakadický pohyb bývá označován jako glisáda. Běžně oko provede 3–4 sakády za sekundu, což znamená, že za den jich vykoná 200 000. Člověk, který se dožije 80 let, tedy za svůj život vykoná téměř šest milionů sakád (Bojko, 2013).



Obr. 4 Znáznornění nejdůležitějších pohybů oka – fixací a sakád. Fixace jsou zobrazeny jako kruhy, jejichž velikost odpovídá délce fixace. Linie reprezentují sakády.

Mezi další pohyby oka patří například smooth pursuit neboli sledovací pohyb, což je pohyb, který oči vykonávají při sledování pomalu se pohybujícího objektu. Příkladem může být například pozorování letadla letícího na obloze. Jestliže se předmět pohybuje rychlostí menší než 20–30°/s, dokáže jej fovea přesně sledovat. Začátek pohybů je opožděn asi o 150 ms. Opoždění za pohybujícím se předmětem je vyrovnáváno korekčními sakádami (Synek a Skorovská, 2014). Sakády a smooth pursuit jsou zcela rozdílné pohyby a řídí je rozdílné části mozku. Na rozdíl od sakád vyžaduje smooth pursuit objekt ke sledování a nelze jej vykonávat ve tmě či před holou zdí. Tabulka 1 znázorňuje typické hodnoty délky trvání, amplitudy a rychlosti nejběžnějších pohybů oka. Hodnoty jsou převzaty z publikace Holmqvista a kol. (2011) a mohou se lišit od hodnot udávaných jinými autory. Například Duchowski (2007) uvádí délku sakád 10–100 ms.

Tab. 1 Typické parametry vybraných pohybů očí (upraveno dle Holmqvist a kol., 2011)

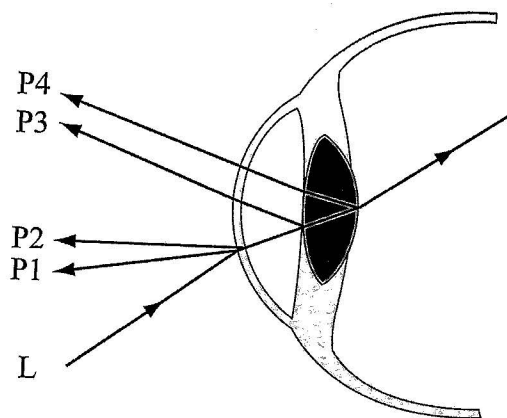
Pohyb oka	Délka (ms)	Amplituda	Rychlost
Fixace	200–300	–	–
Sakáda	30–80	4–20°	30–500°/s
Glisáda	10–40	0,5–2°	20–140°/s
Smooth pursuit	–	–	10–30°/s
Mikrosakáda	10–30	10–40°	15–50°/s
Tremor	–	< 1°	20'/s (max)
Drift	200–1000	1–60'	6–25'/s

V kartografickém výzkumu jsou tedy v naprosté většině případů sledovány pouze fixace a sakády. K jejich identifikaci slouží řada algoritmů (viz kapitola 6.3). Na základě eye-tracking metrik, tedy číselných charakteristik fixací a sakád, se určuje vztah pozorovatele vůči sledovanému obrazu, obtížnost podávané informace, pozornost respondenta a další parametry čtení obrazu (viz kapitola 7.8).

### 1.3 Vývoj metod sledování pohybu očí

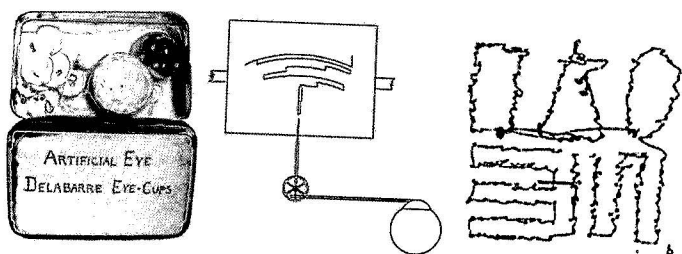
Zájem o lidský zrak a to, jak funguje, je možné pozorovat již ve čtvrtém století před naším letopočtem, kdy si Aristoteles všiml toho, že se obě oči pohybují společně a že například není možné, aby se jedno oko dívalo doleva a druhé doprava. Aristoteles tak došel k závěru, že obě oči je třeba vnímat jako společnou jednotku. O téměř pět set let později se stejnou problematikou zabýval také Ptolemaios, který dokonce prováděl experimentální studie zaměřené na binokulární vidění (Wade, 2010).

O mnoho let později, v devatenáctém století, se problematikou lidského zraku a vidění začal zabývat český fyziolog, anatom a biolog Jan Evangelista Purkyně. Ten je v oblasti sledování pohybu očí velmi známý zejména díky svým popisům obrazů odrážejících se od povrchu oka (tzv. Purkyňovy obrázky / Purkinje images), které mohou být použity pro přesný záznam polohy oka. Existují čtyři viditelné Purkyňovy obrázky (obr. 5). První Purkyňův obrázek (P1) představuje odraz od vnější strany rohovky. Druhý (P2) je odrazem od její vnitřní strany. Třetí (P3) je odrazem od přední strany čočky a poslední (P4) je odrazem od její zadní strany (Glenstrup a Engell-Nielsen, 1995).



Obr. 5 Čtyři Purkyňovy obrázky, odrazy světelného paprsku (L) na různých částech oka: P1 od vnější strany rohovky, P2 od vnitřní strany rohovky, P3 od přední strany čočky a P4 od zadní strany rohovky

To, že se oči nepohybují plynule, popsal poprvé francouzský vědec Émile Javal, který založil a vedl oftalmologickou laboratoř na pařížské Sorbonně. Na konci 19. století popsal takzvané sakády, které pojmenoval podle rychlého pohybu koně při drezúře. Javal ve své studii (1878) rovněž sestrojil zařízení pro záznam pohybu očí, které využívalo odraz zrcátka připevněného na oko zaznamenávaný na fotografickou desku. Na Javala navázal Delabarre (1898), jenž místo zrcátka na oko připevnil sádrovou misku, ke které byl připevněn drátek (obr. 6). Oční pohyby byly poté zaznamenávány pomocí kymografu. Delabarre ve své práci uvádí, že toto zařízení bylo velice přesné při pomalých pohybech očí. Při (přirozenějších) rychlejších pohybech však docházelo k problémům a oční pohyby se nedařilo interpretovat. Jelikož byla tato procedura bolestivá, používal Delabarre ke znečtivění oka roztok kokainu.



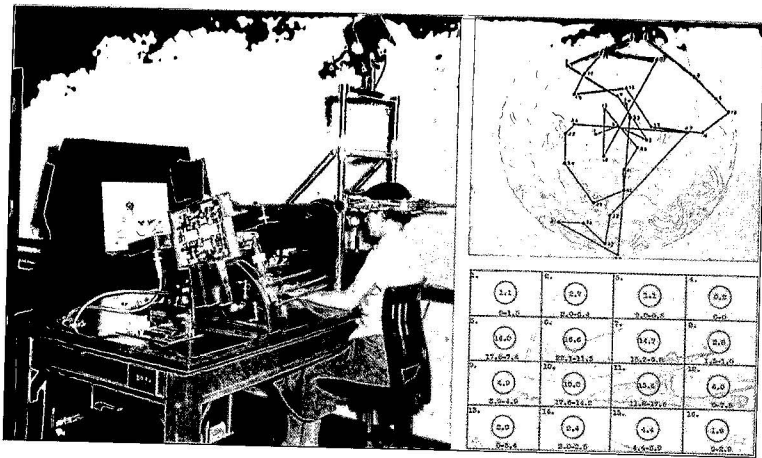
Obr. 6 Sádrové misky (vlevo), ke kterým Delabarre (1898) připevnil drát a pomocí systému převodů (uprostřed) získal záznam pohybu očí (vpravo).

S počátkem 20. století došlo k rychlému nárůstu technologií pro sledování pohybu očí. Pákový přístroj Delabarra byl omezený jeho mechanickými vlastnostmi. Z toho důvodu vznikly nové metody, u kterých nebylo nutné přímé připojení oka a záznamového povrchu. První experimenty s těmito primitivními eye-trackery byly zaměřeny na oblast čtení textu či jednoduchých geometrických obrazců.

Dodge a Cline (1901) později vyvinuli fotografický přístroj, který nevyžadoval žádné uchycení na oko a byl pro účastníky výrazně pohodlnější. Po vzniku tohoto zařízení následovala revoluce ve výzkumu pohybů očí a nárůst počtu experimentů v této oblasti. Dodge a Cline ovšem nebyli jediní a nárůst počtu zařízení pro sledování pohybu očí na začátku 20. století šel ruku v ruce s rychlým vývojem teorií o vazbě mezi pohyby očí (fixace a sakády) a percepce a kognicí.

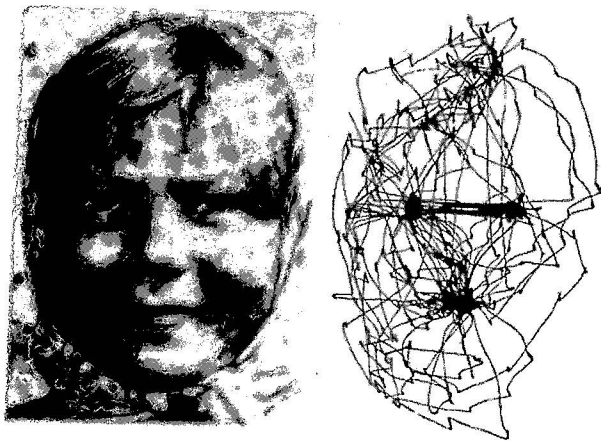
Tyto nové technologie umožnily zjistit, na která konkrétní místa se testovaný subjekt zaměřoval, což později ve 20. století zaujalo bezkonkurenční význam jak v oblasti eye-trackingu, tak ve výzkumu pozornosti. Klíčovou osobností při vzniku této oblasti výzkumu byl Guy Buswell, působící na univerzitě v Chicagu. Jeho nejvýznamnější prací byla monografie „How People Look at Pictures“ (Buswell, 1935). Obsahuje údaje o pohybu očí zaznamenané od celkem 200 respondentů, přičemž každý z nich prohlížel více obrázků. Celkem tedy zaznamenaná data obsahovala více než 2000 záznamů pohybu očí a byla zaznamenána na 5,5 km filmového pásu. Takto rozsáhlý objem naměřených dat je působivý i z hlediska současných standardů, natož při práci s tehdejší technologií, kdy bylo nutné manuálně transformovat horizontální a vertikální polohu oka. Tato práce byla první, která systematicky zkoumala pohyby očí respondentů při prohlížení složitějších stimulů, než jsou texty či jednoduché geometrické obrazce, a představovala něco revolučního v oblasti eye-trackingu (obr. 7). Buswellova monografie obsahuje širokou škálu metod analýzy dat naměřených při prohlížení obrázků, včetně některých překvapivě moderních. Buswell například studoval celkovou distribuci fixací na obrázcích; porovnával několik prvních fixací na obraze s několika posledními; porovnával různé respondenty při sledování stejného stimulu či se zaměřil na vliv zadaných úkolů na zaznamenanou trajektorii. Součástí monografie jsou dokonce i první „heat mapy“, tedy mapy hustoty fixací v jednotlivých částech obrazů, a vymezení „oblastí zájmu“ (Wade, 2010). Buswellovu monografii je tedy právem možné považovat za základ dnešního eye-tracking výzkumu.

Podobnou problematikou jako Buswell se v Moskvě zabýval Alfred Yarbus ve svém výzkumu v padesátých a šedesátých letech 20. století (přeloženo do angličtiny jako Yarbus, 1967). V této práci Yarbus provedl experiment, kdy stejný obraz zobrazil respondentům sedmkrát za sebou, ale pokaždé doplněný jinou otázkou.



Obr. 7 Eye-tracking zařízení použité v roce 1935 Buswellem (vlevo) a výstupy z jeho měření (vpravo)

Tento experiment potvrdil Buswellovo dřívější pozorování, že úkoly dané respondentovi mohou radikálně změnit místa, na která respondent fixuje. Yarbusova práce se stala klasickým dílem v oblasti eye-trackingu a je často uváděna jako jednoznačná demonstrace toho, že „high-level“ faktory (dané úkolem) mohou zastínit jakékoliv „low-level“ faktory (dané stimulem). Kromě toho se poprvé ukázalo, že při pozorování lidského obličeje dochází k fixacím na oči a pusy a ke vzniku trojúhelníku, jak je to patrné z obrázku 8 (Wade, 2010).



Obr. 8 Zobrazení trajektorie pohybu oka zaznamenané během prohlížení obrazu „Dívka od Volhy“. Jasně patrné je umístění fixací v oblasti očí a pusy tvořící trojúhelník (Yarbus, 1967).

Objektivní eye-trackery vyvinuté koncem 19. a začátkem 20. století umožnily zjištění zásadních nových poznatků o skutečné povaze pohybu očí. Pohyby očí nebyly tak hladké a nepřetržité, jak se subjektivně zdálo. Technologický pokrok umožnil řešit nové otázky a identifikoval nové a nečekané otázky v psychologii a fyziologii pohybu očí a jejich vztahu ke kognici.

V 70. letech 20. století se většina prací zaměřila na technická vylepšení eye-trackerů, která měla za cíl především zvýšení jejich přesnosti (Mohamed a kol., 2007). Vzniklo několik velice přesných technik měření pohybu očí, přičemž například metoda Dual Purkinje Image Tracker nebyla doposud z hlediska přesnosti měření překonána.

Obecně existují dva typy technik sledování pohybu oka: ty, které měří polohu oka vzhledem k hlavě, a ty, které měří orientaci oka v prostoru, neboli „Point of Regard“ (Young a Sheena, 1975).

Duchowski (2007) definuje čtyři následující techniky sledování pohybu očí:

- Elektro-OculoGraphy (EOG),
- Scleral contact lens (kontaktní čočky),
- Photo-OculoGraphy (POG) a Video-OculoGraphy (VOG),
- Pupil and Corneal Reflexion Tracking.

Princip **elektrookulografie** (EOG) spočívá v měření rozdílů elektrického odporu kůže pomocí elektrod umístěných kolem očí. Zaznamenáváním poměrně malých rozdílů v kožním odporu byl vypočítán pohyb očí (obr. 9). Jednou z výhod EOG je možnost měřit i pohyb zavřených očí. Toho se využívá například v klinických studiích zaměřených na oční pohyby během spánku. Tato technika měří polohu očí vůči hlavě, proto ji nelze využít ke zjišťování směru pohledu (Point of Regard), pokud není současně měřena pozice hlavy (pomocí head-trackeru). V polovině sedmdesátých let 20. století bylo EOG nejpoužívanější technikou eye-trackingu (Young a Sheena, 1975), ale ze čtyř výše zmíněných technik je nejméně přesná (Yarbus a kol., 1967).



Obr. 9 Ukázka okulografie využívající elektrod kolem očí (Larson a kol., 2017)

Naopak jednou z nejpřesnějších technik měření pohybu očí je použití mechanického nebo optického referenčního objektu namontovaného na **kontaktní čočku**, který je pak umístěn přímo na oku. První zařízení tohoto typu (využívané například Delabarrem) využívala sádrové misky fyzicky propojené se záznamovým perem například pomocí drátu. Tato technika se později vyvinula a začaly být používány moderní kontaktní čočky umístěné na rohovce a bělimě. Na tuto čočku byla umístěna řada optických či mechanických zařízení, například obsahujících fosfor nebo drátěnou cívku. Poloha a orientace této cívky (a tím pádem i oka) se vypočítala pomocí měření změn proudů indukovaných třemi navzájem kolmými zdroji magnetického pole (obr. 10) (Hammoud a Mulligan, 2008). Vložení čočky do oka vyžaduje péči a cvičení a její použití je pro respondenta nepohodlné. I tato technika měří polohu oka vzhledem k hlavě.



Obr. 10 Kontaktní čočka obsahující drátěnou cívku (vlevo) a externí rám pro měření změn magnetického pole (vpravo)<sup>2</sup>

Další z technik měření pohybu očí je **foto- nebo videookulografie**, jež spočívají v měření pozice rozlišitelných částí oka. Používá se zejména detekce tvaru zornice, pozice limbu (hranice mezi rohovkou a bělimou) a korneálního odrazu (odrazu blízko umístěného (infračerveného) světelného zdroje od rohovky).

Interpretace pohybů očí poskytovaných těmito technikami může nebo nemusí být prováděna automaticky a může zahrnovat i prostou vizuální kontrolu očních pohybů (obvykle z videozáznamu). Manuální procházení videozáznamů je ale extrémně zdlouhavé, náchylné k chybám a je omezeno vzorkovací frekvencí videozariadení. Některé z těchto metod vyžadují, aby byla hlava pevně upevněna pomocí opěrky hlavy nebo pomocí zubů (pomocí tzv. bite-baru) (Young a Sheena, 1975).

<sup>2</sup> [http://work.lhaslwanter.at/Kinematics/html/03\\_RotMats.html](http://work.lhaslwanter.at/Kinematics/html/03_RotMats.html)

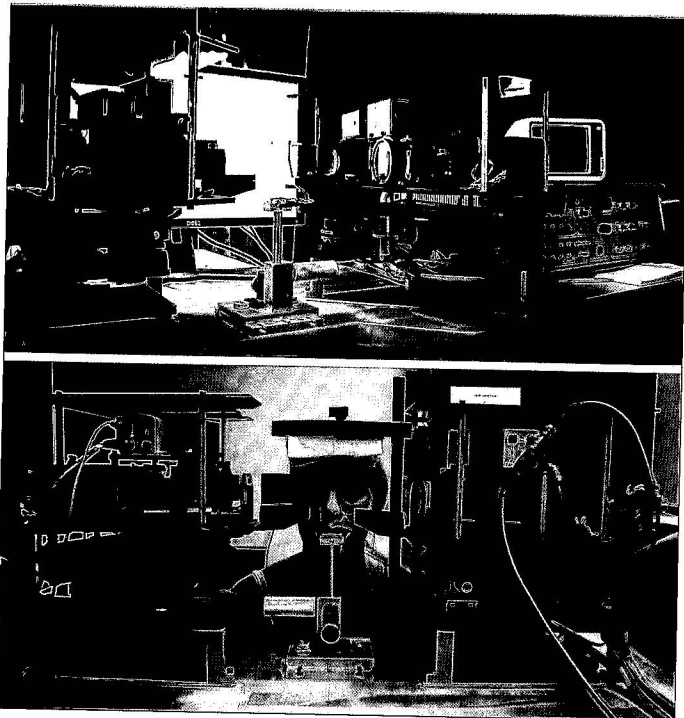
Automatické sledování limbu často zahrnuje použití infračervených fotodiod namontovaných na rámu okolo oka (obr. 11). Výhoda sledování limbu spočívá v tom, že na rozdíl od zornice není velikost limbu ovlivněna různými světelnými podmínkami. Naopak nevýhodou je časté překrytí limbu očními víčky.



Obr. 11 Limbus tracker<sup>3</sup>

Z hlediska prostorové přesnosti i frekvence záznamu je dosud nepřekonanou technikou měření dvou Purkyňových obrázků využívané zařízením **Dual Purkinje Image (DPI) tracker**. Toto zařízení (Crane a Steele, 1985) využívá prvního a čtvrtého Purkyňova obrázku, tedy odrazů světelného paprsku od vnější strany rohovky a vnitřní strany čočky. Ke sledování těchto dvou odrazů používá zařízení systému zrcadel a velmi rychlých optických serv. Problematické je zejména to, že čtvrtý Purkyňův obrázek je asi 400× slabší než první. Díky své vysoké přesnosti se DPI tracker využívá jako referenční při porovnávání přesnosti eye-trackerů. I při používání DPI trackeru je však nezbytná fixace hlavy a v ideálním případě i zubů. Dual Purkinje Image tracker od společnosti Fourward Technologies je znázorněn na obrázku 12. V dolní části je poté znázorněna fixace hlavy.

<sup>3</sup> <http://slideplayer.com/slide/5952194/>



Obr. 12 Dual Purkinje Image tracker (nahore)  
a nezbytná fixace hlavy při jeho používání (dole)<sup>4</sup>

Ačkoliv jsou výše uvedené techniky obecně vhodné pro měření pohybů očí, neposkytují Point of Regard, tedy směr pohledu. Pro jeho zjištění musí být hlava zafixována tak, aby se shodovala poloha oka vzhledem k hlavě se směrem pohledu.

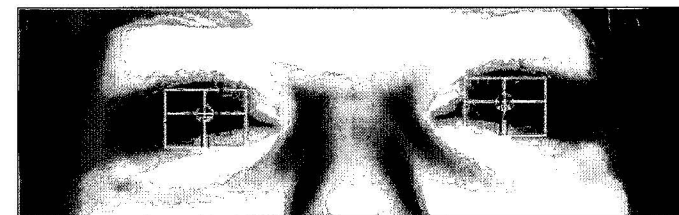
Pro zjištění směru pohledu je nezbytné měření několika charakteristik očí zároveň tak, aby se odlišil pohyb hlavy od otáčení očí. Takovými charakteristikami očí jsou nejčastěji pozice středu zornice a infračerveného odrazu od rohovky. Tyto dvě charakteristiky využívá v dnešní době nejčastěji aplikovaná technika sledování pohybu očí **Pupil and Corneal Reflexion Tracking**. Její princip spočívá v detekci středu zornice a korneálního (kornea = rohovka) odrazu přímého paprsku infračerveného světla. Jak už bylo uvedeno výše, existují celkem čtyři Purkyňovy obrázky. Korneální odraz je pouze jiné označení pro první Purkyňův obrázek, který je nejčastěji využíván pro měření pozice pohledu. Vzájemná poloha středu zornice a korneálního odrazu se při otáčení očí mění, ale při menších pohybech hlavy zůstává relativně konstantní.

<sup>4</sup> <http://twiki.cis.rit.edu/twiki/bin/view/MVRL/DualPurkinjeImageTracker>



Obr. 13 Eye-tracker umístěný pod monitorem<sup>5</sup>

Eye-tracker je nejčastěji umístěn pod monitorem, na kterém je zobrazován studovaný obraz neboli stimulus (obr. 13). Součástí tohoto zařízení je jedno nebo více infračervených světel, které svítí směrem na uživatele. Zařízení také obsahuje kameru, která snímá oči respondenta. Ta na základě rozpoznávání obrazu nalezne střed zornice a korneální odraz infračerveného světla (obr. 14). Ze vzájemné polohy těchto dvou bodů zařízení vypočítá směr pohledu (Point of Regard).

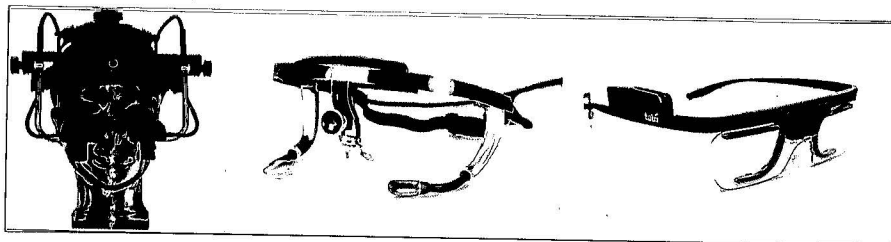


Obr. 14 Snímek z kamery umístěné na eye-trackeru. Bílým křížem je označen střed zornice, černým korneální odraz.

Mnoho informací o základních mechanismech lze shromáždit z reakcí v kontrolovaných laboratorních studiích, ale již od počátku eye-tracking výzkumu existoval velký zájem o studium chování respondentů v reálném prostředí. Dnes již existuje mnoho mobilních systémů umožňujících sledování pohybu očí přímo v reálném prostředí. Tyto eye-trackery rovněž využívají Pupil and Corneal Reflexion tracking a respondent je má umístěné přímo na hlavě jako brýle nebo jsou eye-trackery namontované na nejrůznější přilby (obr. 15).

<sup>5</sup> <http://eyegaze.com/wp-content/uploads/Study-active-checking-more.jpg>

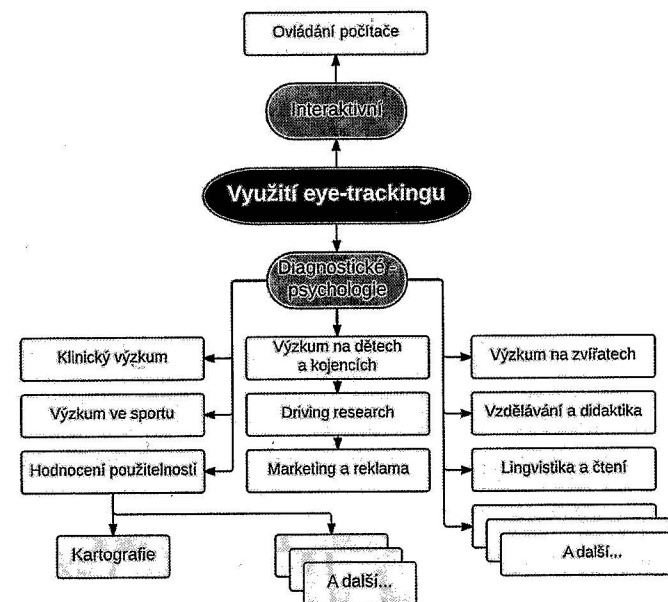




Obr. 15 Mobilní eye-trackery od společností SR research (vlevo), Ergoneers (uprostřed) a Tobii (vpravo)

## 2 VYUŽITÍ EYE-TRACKINGU

Eye-tracking je využíván v mnoha oblastech lidské činnosti. V této publikaci není možné obsáhnout celé spektrum využití sledování pohybu očí, a ani to není jejím cílem. Tato kapitola tedy slouží pouze jako velmi neúplný přehled širšího využití eye-trackingu. Duchowski (2007) rozděluje využití eye-trackingu do dvou základních kategorií. Jedná se o interaktivní využití, kdy eye-tracker slouží k ovládnutí počítače, a diagnostické využití, kdy jsou pohyby očí zaznamenávány pro odhalení toho, které části stimulů upoutaly respondentovu pozornost. Oblasti využití eye-trackingu popsané v této knize jsou znázorněny na obrázku 16.



Obr. 16 Schéma využití eye-trackingu

## 2.1 Interaktivní využití eye-trackingu

Eye-tracking lze využít pro ovládání počítače. Pohled tak nahrazuje nebo doplňuje běžné periferie, jako je klávesnice nebo myš. S eye-trackerem dokonce experimentoval i známý teoretický fyzik Stephen Hawking.<sup>6</sup> Eye-tracking jako doplněk běžných periférií je již dnes instalován do výkonných herních notebooků.<sup>7</sup>

Eye-tracker však může klávesnici a myš úplně nahradit. Tento přístup najde své využití především u kvadruplegiků, kteří nemají možnost jiným způsobem počítač ovládat. Nevýhodou tohoto přístupu je zejména to, že lidé běžně využívají zrak pro získání informací o okolí. To způsobuje problémy v případě, že má zrak sloužit jak pro získání informací, tak pro ovládání. Tento problém bývá označován jako Midasův dotyk dle starověké pověsti o králi Midasovi, který proměnil ve zlato vše, čeho se dotkl. Uživatel se například soustředí (fixuje) na určité místo, protože ho zaujalo, a chce si jej prohlédnout. Systém ale tuto fixaci vyhodnotí jako interakci (klik).

Zaměření se na určitý objekt pomocí zraku je velmi jednoduché, intuitivní a není třeba žádného tréninku. Nicméně v případě postižených lidí toto nemusí platit. Těžce postižení od narození nejsou zvyklí cokoli fyzicky ovládat. Prosté naučení se ovládání kurzoru pomocí pohledu tak může těmto uživatelům trvat od několika minut až po několik měsíců (Gips a kol., 1996).

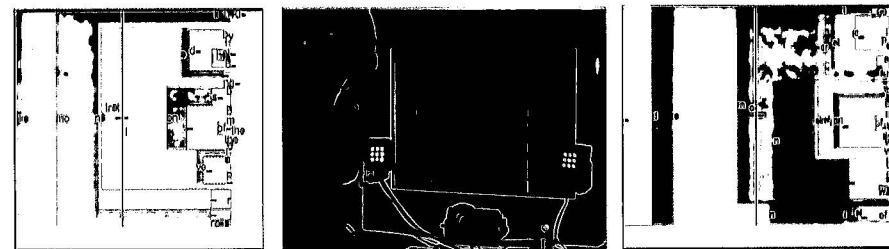
Většina eye-tracking systémů určených pro psaní textu využívá virtuální klávesnici umístěnou na obrazovce. Psaní na takové virtuální klávesnici se skládá ze tří fází (Majaranta a Riihã, 2002). V první z nich se uživatel zaměří na určité písmeno (focus). Druhou fází je odezva (feedback), která je pro postižené uživatele velmi důležitá. Protože v dosavadním životě nebyli schopni fyzicky manipulovat s objekty, je nutné, aby se vybrané písmeno vizuálně označilo, případně bylo toto označení doprovázeno akustickým signálem. Ve třetí fázi dochází k potvrzení výběru (selection), a tím pádem k interakci, tedy napsání daného písmena. Tato selekce nejčastěji probíhá pomocí Dwell time, tedy fixací na dané klávese o předem stanovené délce (většinou 600–1000 ms). Kromě Dwell time je možné využívat například mrknutí (Rasmusson a kol., 1999), stisknutí tlačítka, pohyb hlavy (Beukelman a Mirenda, 2005) či předem nastavená gesta očí (například pohled doleva).

Rozložení virtuálních klávesnic je ve většině případů odlišné od klasického QWERTY. Uživatelé se například mohou lépe orientovat v klávesnici seřazené dle abecedy. Ward a MacKey (2002) vytvořili nástroj Dasher,<sup>8</sup> sloužící k psaní textu pohledem. Namísto klasického rozložení kláves jsou jednotlivá písmena zobrazována tak, jak je znázorněno na obrázku 17.

6 <http://www.hawking.org.uk/the-computer.html>

7 <https://tobiigaming.com/alienware17-eye-tracking-how-to/>

8 <http://www.inference.org.uk/dasher/>



Obr. 17 Prostředí aplikace Dasher určené pro psaní textu pomocí pohybu očí

Po pohledu na určité písmeno program vyhodnotí nejpravděpodobnější písmena, která by mohla následovat. Ta se zvýrazní tak, aby je uživatel rychleji viděl. Autoři uvádějí, že po hodině tréninku jsou uživatelé schopni psát rychlostí až 25 slov za minutu. Rychlost psaní všemi deseti na běžné klávesnici přítom je asi 75 slov za minutu.

Další volně dostupnou aplikací pro ovládání počítače pomocí pohybu očí je GazeSpeaker.<sup>9</sup> Ten je kompatibilní se zařízením EyeTribe a funguje na principu předem definovaných karet se slovy. Každé zvolené slovo se automaticky napíše a zároveň dojde k jeho přečtení. GazeSpeaker tak mohou využívat i postižení, kteří nemohou mluvit. Kombinace GazeSpeakeru a EyeTribe se využívá například k popularizačním účelům v olomoucké Pevnosti poznání.

## 2.2 Diagnostické využití eye-trackingu

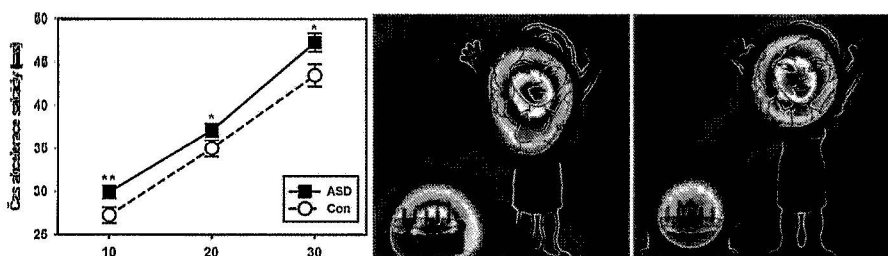
Jednou z nejširších a také nejstarších oblastí, ve které eye-tracking našel své využití, je psychologie. Psychologický eye-tracking výzkum je možné rozdělit do tří základních oblastí. Jedná se o neuropsychologii, rozvojovou psychologii a kognitivní psychologii. Psychologie pokrývá velkou část oblastí, kde je eye-tracking využíván, a je proto možné ji považovat za jakousi nadkategorii. Pod neuropsychologii by tak mohl patřit níže zmíněný klinický výzkum. Rozvojová psychologie zkoumá vývoj dětí či zvířat. Stejně tak je možné sledovat vývoj při čtení či učení. Nejširší oblastí je pak kognitivní psychologie, pod kterou mohou spadat takřka všechna využití eye-trackingu, a to včetně výše zmíněných. Obsáhlou publikaci zaměřenou pouze na kognitivní psychologii napsala Matlin (2014). Z důvodu provázanosti s ostatními oblastmi výzkumu byla psychologie vyčleněna na začátek této kapitoly a popisovány budou jednotlivé podkategorie.

9 <http://www.gazespeaker.org/>

### 2.2.1 Klinický výzkum

V klinickém výzkumu slouží eye-tracking zejména při identifikaci onemocnění zraku, ale také při sledování a diagnostice duševních a nervových onemocnění, jako je porucha autistického spektra, ADHD či Parkinsonova choroba.

Velmi časté jsou studie zaměřené na autismus. Jedinci s poruchami autistického spektra (ASD) vykazují atypické trajektorie pohybu očí zejména během společenské interakce a při prohlížení tváří, ale i při sledování jednodušších a nesociálních stimulů (Schmitt a kol., 2014). Eye-tracking výzkum autistických pacientů může být zaměřen na analýzu sakadických pohybů očí (např. Schmitt a kol., 2014, obr. 18 vlevo) nebo na prostorové umístění fixací. Toho se využívá zejména při použití stimulů zobrazujících lidskou tvář. Příkladem takové studie může být například magisterská práce Kurečkové (2016) z Masarykovy univerzity v Brně. Jako stimuly sloužily obrázky lidských postav a respondenty byli žáci speciální základní školy pro děti s autismem. Data byla měřena pomocí zařízení EyeTribe a aplikace OGAMA (obr. 18 vpravo).



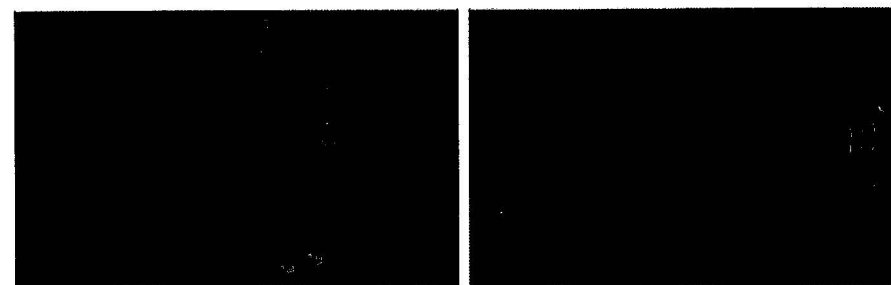
Obr. 18 Ukázka výzkumu pacientů s autismem.

Vlevo je analýza průběhu sakád (Schmitt a kol., 2014), vpravo analýza rozmístění fixací na stimulu s postavou (Kurečková, 2016).

Velmi zajímavou problematikou, kdy je eye-tracking využíván v klinickém výzkumu, je vyšetření pacientů s neglect syndromem po mozkové mrtvici. Neglect syndrom je porucha pozornosti, vnímání a orientace v kontralaterální polovině prostoru bez porušení primárních motorických a senzorických funkcí (Ambler, 2011). Zjednodušeně řečeno, jedná se o poruchu, při které pacienti nevnímají podněty z jedné strany. V akutním stadiu, zejména po pravostranných lézích, bývá neglect výrazný. Pacientům se nedaří obléci polovinu svého těla, oholit či nalíčit polovinu tváře, jedí pouze z jedné poloviny talíře, a stěžují si proto na malé porce (Hájková, 2016).

Na oddělení rehabilitace Fakultní nemocnice v Olomouci se pro vyšetření pacientů s neglect syndromem využívá zařízení EyeTribe a aplikace OGAMA. Touto problematikou se zabývala například Hájková (2016), která

v prezentovaných stimulech vytvořila vertikální oblasti zájmu a hodnotila, kolik času v nich pacienti stráví. Dále data vizualizovala pomocí scanpath a attention map (obrázek 19).



Obr. 19 Vizualizace eye-tracking dat pacienta s levostrannou hemiparézou (Hájková, 2016)

Záznam pohybu očí je možné kombinovat se záznamem z dalších psychofyziologických zařízení (viz kapitola 3.4) Velmi často rovněž dochází ke kombinaci s funkční magnetickou rezonancí (fMRI).

### 2.2.2 Výzkum na dětech a kojencích

Další oblastí je výzkum prováděný na dětech a kojencích. Zde se může opět jednat o výzkum spojený s autismem, ale také s vývojovou psychologií. Výzkum se provádí zejména za účelem pochopení toho, kam se kojenci a děti dívají při sledování lidí provádějících různé úkoly, kdy začnou rozpoznávat své rodiče a jak dekodují záměry sledovaných osob. Několikanásobná prezentace stejných stimulů v průběhu vývoje dítěte umožňuje zkoumání časového průběhu učení a také toho, jak se postupně mění pozornost kojenců.

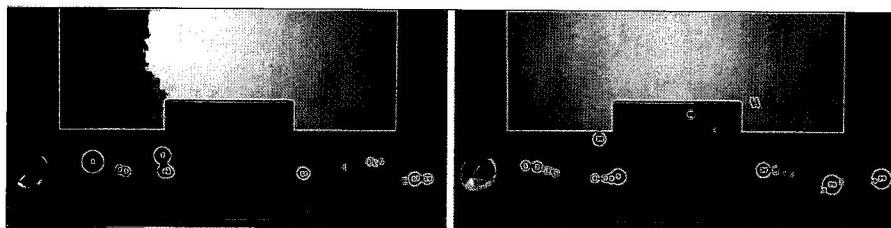
Při záznamu pohybu očí u dětí je nutné upravit proces kalibrace. Místo pohybující se tečky jsou proto většinou používány animované obrázky hraček, na které se dítě lépe soustředí. Kalibrace bývá rovněž často doplněna zvukem.

Schopnost pamatovat si dočasně zakryté objekty je důležitým mezníkem v raném dětství. Tato schopnost pomáhá dítěti hledat souvislosti i mezi věcmi, které nejsou v daném okamžiku viditelné, předpovídat budoucí události a úspěšně komunikovat s okolím. Zajímavý experiment prováděli Kochukhova a Gredebäck (2007), když zjišťovali, jak se tato schopnost vyvíjí během prvního roku života. Kojencům byly zobrazeny videozáznamy, v nichž se kutálel míč ve scéně zobrazené na obrázku 21. Od čtvrtého měsíce života dítě sleduje míč až do okamžiku, kdy zmizí. Jakmile se opět objeví, dítě pokračuje ve sledování. Od šestého měsíce už je dítě schopné si představit,

že se míč pohybuje i za deskou, protože se na místo, kde se míč objeví, dívá ještě dřív, než se tak skutečně stane.



Obr. 20 Ukázka eye-tracking výzkumu prováděného na kojencích. Vlevo využití mobilního eye-trackeru (Franchak a kol., 2010), vpravo využití eye-trackeru pod monitorem<sup>10</sup>



Obr. 21 Experiment zkoumající představivost kojenců. Na obrázku vlevo je zobrazena trajektorie pohybu oka pětiměsíčního kojence. Na obrázku vpravo pak kojence sedmiměsíčního.<sup>11</sup>

### 2.2.3 Výzkum na zvířatech

Valná většina eye-tracking experimentů je samozřejmě prováděna na lidech. Přesto je možné sledovat pohyby očí i u zvířat (obr. 22). Studie jsou často prováděny na primátech (např. Kano a Tomonaga, 2009; Simpson a kol., 2016), ale také na kočkách (Körding a kol., 2001), psech (Somppi a kol., 2012) či pávech (Yorzinski a kol., 2017).

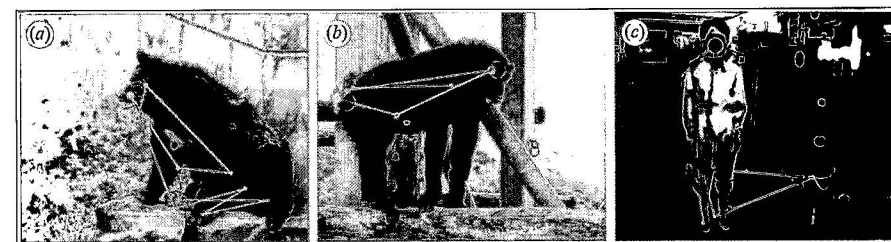
<sup>10</sup> <https://www.tobiiipro.com/fields-of-use/infant-child-research/developmental-psychology/>

<sup>11</sup> <https://www.tobiiipro.com/fields-of-use/infant-child-research/customer-cases/uppsala-university/>



Obr. 22 Různé typy eye-trackerů používaných při nahrávání pohybu očí zvířat<sup>12</sup>

Kano a Tomonaga (2009) realizovali eye-tracking studii na šimpanzech. Jako stimuly sloužily obrázky lidských a šimpanzích postav. Ukázalo se, že pohyby očí u šimpanzů a lidí jsou velmi podobné. Výrazně vyšší počet fixací byl zaznamenán na zobrazené postavě než na okolí. Více fixací bylo zaznamenáno na hlavě než na zbytku postavy. Na obrázku 23 jsou zobrazeny trajektorie pohybu očí člověka (červená) a šimpanze (žlutá).



Obr. 23 Porovnání trajektorií pohybu očí člověka (červená) a šimpanze (žlutá) (Kano a Tomonaga, 2009)

### 2.2.4 Výzkum ve sportu

Eye-tracking se využívá i ve sportu, kde slouží k odhalení problémů souvisejících se ztrátou pozornosti, špatným odhadem trajektorie, strategií vizuálního vyhledávání či koordinací ruky a oka. Často je cílem studií nalézt rozdíly mezi profesionálními sportovci a amatéry (obr. 24).

Komplexní přehled eye-tracking studií ve sportu sestavili Kredel a kol. (2017). Ve svém článku shrnují čtyřicet let výzkumu a popisují téměř 60 studií. V oblasti sportu je klíčové dodržet ekologickou validitu experimentu – tedy to, aby probíhal v co nejrealnějších podmínkách. Toho je dosaženo použitím mobilních eye-tracking zařízení. Z 60 popisovaných

<sup>12</sup> <https://www.tobiiipro.com/fields-of-use/psychology-and-neuroscience/primate-canine/>  
<http://visagetechnologies.com/uploads/2015/10/blog-dog-gaze.jpg>  
<https://www.theatlantic.com/photo/2012/12/animals-in-the-news/100430/#img26>  
<http://www.vision-systems.com/articles/2013/08/eye-tracking-cameras-study-impact-of-peacock-tail-display-on-female-birds.html>

studii jich 51 využívalo právě mobilní zařízení. Přesto pouze 40 % z nich bylo testováno v reálných podmínkách, zatímco zbývajících 60 % bylo prováděno v laboratoři. Z přehledu studií obsažených v článku Kredela a kol. (2017) je patrné, že eye-tracking se nejčastěji využívá zejména v míčových sportech, jako je basketbal, fotbal či tenis, squash či badminton, ale také ve sportech, jako je lední hokej, karate, cyklistika nebo box. Jak se dalo očekávat, valná většina studií byla zaměřena na sporty, ve kterých je zrak nutný nejen pro kontrolu pohybu, ale také pro rozhodování.



Obr. 24 Využití mobilních eye-trackerů ve sportu<sup>13</sup>

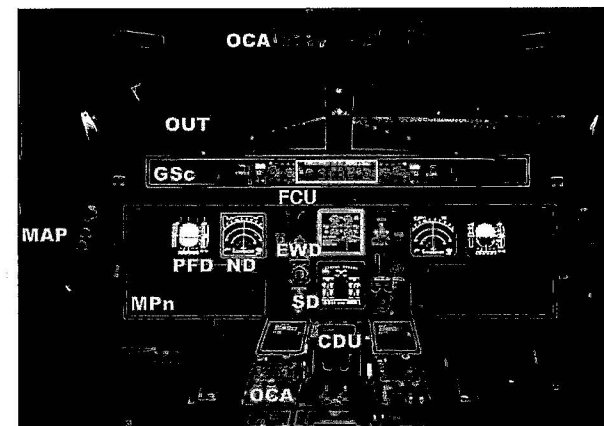
### 2.2.5 Využití v dopravě

Eye-tracking nabízí jedinečný vhled do chování lidí při různých činnostech. Zkoumání interakce člověka a prostředí je velmi důležité i v kontextu bezpečnosti dopravy. Tradiční metody měření lidské výkonnosti často zahrnují měření reakční doby a přesnosti, například jak rychle osoba dokončí úkol, jak dobře je tento úkol vykonáván atd. Stejně důležitá je však i analýza jednotlivých kroků, které vedly ke splnění úkolu. Sledování pohybu očí je v tomto kontextu velmi důležité, protože umožňuje vhled do vizuálních a kognitivních aspektů výkonu či sledování pozornosti respondenta. Z toho důvodu je eye-tracking často využíván pro testování řidičů či pilotů.

V letectví začaly být eye-trackery využívány už v roce 1947, kdy byly hodnoceny pohyby očí armádních pilotů při přistávání. Využívalo se přitom analýzy videozáznamu očí jednotlivých pilotů (Fitts a kol., 1950). S rozšířením leteckých simulátorů začal být eye-tracking aplikován při analýze pohybů očí pilotů během simulovaných letů. Příkladem může být studie Anderse (2001), který zaznamenával pohyby očí a hlavy profesionálních pilotů a způsob jejich ovládní letadla, včetně povědomí o rozložení kokpitu. Kromě pilotů je často testován také výkon a reakce letových dispečerů (např. Ellis, 2009), případně rozložení displayů a jejich obsahu, včetně map (obr. 25). Pro záznam pohybu očí v komplexním prostředí, jako je kokpit letadla či simulátoru, lze buď využít mobilní eye-tracker nebo jsou dostupné

<sup>13</sup> <http://www.4view.com/new-page/>  
<https://www.tobiipro.com/fields-of-use/professional-performance/customer-cases/swiss-ice-hockey-association/>  
<http://www.skysports.com/watch/video/sports/f1/10319019/eye-tracking-with-nico-hulkenberg>

sofistikované eye-trackery složené z několika kamer, jejichž záznam se následně spojí do jednoho, který obsáhne celý kokpit. Příkladem takové technologie je eye-tracker od společnosti SmartEye,<sup>14</sup> který umožňuje najednou integrovat až 8 kamer (eye-trackerů), a obsáhnout tak 360° zorné pole. V Česku se výzkumem využívajícím eye-tracking v leteckém průmyslu zabývá společnost Honeywell v Brně.



Obr. 25 Rozmístění oblastí zájmu okolo jednotlivých přístrojů v kokpitu leteckého simulátoru (Duchowski, 2007)

Eye-trackery se využívají i v automobilovém průmyslu, a to opět pro zpětnou analýzu pohybu očí řidičů, ale také například jako systém varování před mikrosránkem (například Blackbird AntiSleep).<sup>15</sup> O četnosti využití eye-trackingu v dopravních studiích svědčí i to, že mezinárodní organizace pro normalizaci (ISO) vydala „pokyny týkající se podmínek a měření týkající se sběru a analýzy údajů o vizuálním chování řidiče“.<sup>16</sup> Jsou zde definovány pojmy jako fixace, sakáda, pohled (glance) atd. Obsáhlé review literatury zabývající se touto problematikou napsali Young a Regan (2007) a Kircher (2007). Studie zaměřené na automobily a řidiče jsou velmi často prováděny v simulátorech. Jejich výhodou je to, že prostředí může být kontrolováno, a všichni respondenti se tak ocitnou ve stejné situaci. Rovněž je možné analyzovat chování řidičů v nebezpečných situacích, což by v běžném provozu z etických důvodů nebylo možné. Velmi často je pomocí eye-trackingu v automobilovém průmyslu sledováno rozptýlení řidiče, a to například při manipulaci s GPS navigací (Chiang a kol., 2004).

<sup>14</sup> <http://smarteye.se/>

<sup>15</sup> <http://smarteye.se/applied-solutions>

<sup>16</sup> <https://www.iso.org/obp/ui/#iso:std:iso:15007:-1:ed-2:vi:en>

V Česku se touto problematikou zabývá Laboratoř interaktivních vozidlových simulátorů ČVUT v Praze,<sup>17</sup> která je vybavena mobilním eye-trackerem od společnosti SMI.

Kromě simulátorů se nabízí využití virtuální reality kombinované s eye-trackingem. Příkladem takové studie z oblasti dopravy je práce Vora a kol. (2002), kteří využili virtuální realitu pro trénink vizuální inspekce letadel.

## 2.2.6 Výzkum ve vzdělávání a didaktice

Ve vzdělávání a didaktice existují dva základní přístupy pro využití eye-trackingu. Jedná se o hodnocení vytvořených materiálů a hodnocení uživatelů.

V prvním případě jsou vytvořené materiály, jako jsou pracovní listy či učebnice, použity jako stimuly, které jsou následně hodnoceny. Jarozka a kol. (2017) ve svém článku uvádějí teoretický rámec a výzkumnou agendu v této oblasti. V článku je popsána ukázková studie zaměřená na hodnocení učebního materiálu. V ní byly porovnávány dvě varianty multimediálního učebního textu z oblasti výtvarného umění. Testováno bylo 16 studentů, pro které byl tento materiál určen, což zajistilo ekologickou validitu experimentu. V první verzi materiálu byly doplňující obrázky zobrazeny na jedné straně materiálu, zatímco text byl zobrazen na straně druhé. Studenti tedy museli rozdělit pozornost mezi obrázky a text a vizuálně hledat související informace. V druhé variantě byly obrázky umístěny přímo v textu na místech, kde na ně bylo odkazováno. To odpovídalo klasickému integrovanému designu, který studentům umožňuje zpracovat multimediální informaci přímo. Zajímavé bylo, že studenti dosáhli lepších výsledků při první variantě stimulu. Záznam pohybu očí ukázal, že respondenti doplňující informace téměř nevyužívali. Studenti tedy doplňující informace pro splnění zadaného úkolu nezbytné, a nezdržovali se s nimi.

Eye-tracking je možné rovněž využívat přímo během výuky. První možností je vybavení učitele eye-tracking brýlemi, přičemž jsou hodnoceny jeho reakce na žáky. Alternativou je sledování videozáznamu vyučovací hodiny. Takovou studii provedli například Yamamoto a Imai-Matsumura (2013), kteří 43 učitelům pustili minutový videozáznam z vyučovací hodiny. Dva ze studentů na videu neuposlechli učitelův příkaz a nezavřeli učebnice. Experiment byl zaměřen na to, zda si těchto studentů učitelé všimnou. Druhým způsobem, jak lze eye-tracking využít přímo ve výuce, je vybavení počítačové laboratoře eye-trackery. Učitel tak může přímo pozorovat, zda se studenti soustředí a kam přesně se dívají. Takovou učebnou je vybavena například Fakulta informatiky a informačních technologií Slovenské technické univerzity v Bratislavě (obr. 26).

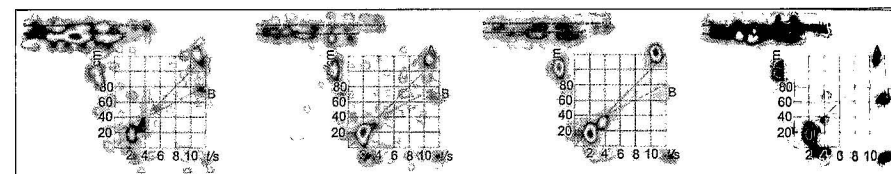
<sup>17</sup> <https://www.fd.cvut.cz/o-fakulte/laborator/41>



Obr. 26 Učebna vybavená dvaceti eye-trackery od společnosti Tobii na Fakultě informatiky a informačních technologií Slovenské technické univerzity v Bratislavě<sup>18</sup>

Uživatelé byli hodnoceni rovněž ve studii van Meeuwena a kol. (2014). Třicet jedna řídicích letového provozu na třech různých úrovních odborných znalostí (nováčky, středně pokročilí, experti) řešilo devět různých úkolů na radarové obrazovce. Kromě pohybů očí byla měřena kognitivní zátěž a správnost řešení. U expertů bylo zjištěno, že používají efektivnější trajektorie pohybu očí, a rovněž správnost řešení u nich byla vyšší.

V Česku se využitím eye-trackingu v didaktice fyziky zabývá například Martina Kekule z Univerzity Karlovy. Ve své studii (Kekule, 2015) uvádí, že typickým výzkumným námětem je srovnání strategií žáků, kteří dosahují ve fyzice dobrých výsledků, s žáky podprůměrnými. Na obrázku 27 je pomocí attention map znázorněno řešení zadaného úkolu nad kinematickým grafem čtyřmi skupinami respondentů.



Obr. 27 Attention mapy pro čtyři skupiny respondentů lišící se úspěšností v testu. Úspěšnost klesá zleva doprava (Kekule, 2015)

Na rozdíl od kartografie je v didaktice žádoucí, aby bylo zadání úkolu obsaženo přímo ve stimulu. Je tak možné analyzovat, kolik času respondenti strávili čtením zadání, jak často se k němu vraceli atd.

<sup>18</sup> <http://ux.fit.stuba.sk/>

Porovnání strategií expertů a laiků se často využívá například v medicíně. Na základě ideálního průběhu trajektorií pohybu oka zkušených expertů se mohou medicí učít číst snímky z lékařských přístrojů. Příkladem může být analýza histologických snímků (Kahl a kol., 2015).

Tématem využití eye-trackingu ve vzdělávání se zabývá například obsáhlá publikace Was a kol. (2016), ve které je možné najít bližší informace o této problematice.

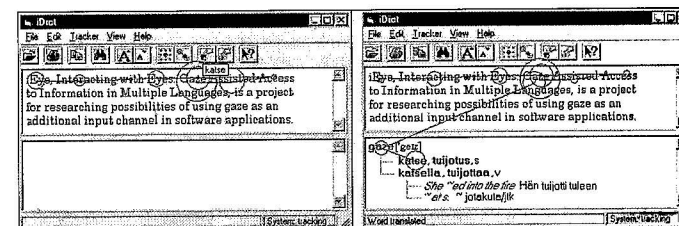
### 2.2.7 Výzkum čtení

Jak už bylo uvedeno výše, první pokusy se záznamem pohybu očí byly testovány na stimulech obsahujících text nebo jednoduché geometrické obrazce. Jako první éra eye-tracking výzkumu při čtení bývá tedy označováno období od roku 1879, kdy s ním experimentoval Javal, až do roku 1920. V tomto období byl výzkum zaměřen především na samotnou podstatu očních pohybů. Podobná situace nastala v takzvané druhé éře výzkumu, přestože tehdy vznikla jedna z klíčových publikací v této oblasti (Tinker, 1946). Ta bývá přirovnávána k Buswellově práci popsané v kapitole 1.3. Ve své další práci Tinker (1958) uvádí, že téměř vše, co mohlo být pomocí sledování pohybu očí o čtení zjištěno (pomocí tehdejší technologie), už zjištěno bylo. Od konce padesátých až do poloviny sedmdesátých let minulého století opravdu téměř žádný eye-tracking výzkum zaměřený na čtení neprobíhal. Třetí éra výzkumu čtení začala právě v polovině sedmdesátých let a vyznačovala se zlepšením systémů pro záznam polohy oka, které umožnily přesnější a snadnější měření (Rayner, 1998).

Ve výzkumu čtení se často pracuje s takzvanými regresemi (regressions) neboli zpětnými sakádami, tedy sakádami v opačném směru, než je směr textu. V češtině je text psaný zleva doprava, a regrese jsou tedy zprava doleva). Rayner (1998) uvádí, že regrese představují asi 10–15 % všech sakád zaznamenaných během čtení textu.

Velká část eye-tracking výzkumu čtení je zaměřena na respondenty trpící dyslexií, kteří provádějí větší počet zpětných sakád. Je prokázáno, že normální čtenáři mají pozornost při zrakovém vnímání rozloženou stupňovitě. Jejich pozornost je tedy nejvíce zaměřena na objekt, který je ve středu zorného pole ve žluté skvrně (makule), a směrem k okrajům sítnice pozornost ke zrakovým stimulům klesá. Zraková pozornost dyslektiků je v prostoru rozložena difúzně a jejich citlivost ke stimulům ve středu zrakového pole je stejná jako k těm, které jsou v jeho periferii (Facoetti a kol., 2003). Toto odlišné rozložení pozornosti se týká zejména pravé poloviny zorného pole, tj. pole důležitého pro plánování dopředných sakád (Facoetti a Molteni, 2001). Pomocí speciálního tréninku lze zrakovou pozornost u dyslektických pacientů normalizovat, a tím dosáhnout významného zlepšení čtení (Facoetti a kol., 2003).

Bylo vyvinuto několik nástrojů využívajících eye-tracking, které čtenářům pomáhají. Příkladem může být aplikace The Reading Assistant (Sibert a kol., 2000), která automaticky rozpozná slovo, kterému má čtenář problém porozumět, a toto slovo se automaticky nahlas přehraje. Podobnou aplikací je iDict (Hyrsykari a kol., 2000), jež pomáhá čtenářům při čtení cizojazyčného textu. Aplikace rovněž sama pozná, kterému slovu čtenář nerozumí, a přímo nad ním se zobrazí překlad z online slovníku. Pokud čtenář slovu stále nerozumí, zobrazí se celý záznam daného slova ze slovníku (obr. 28).



Obr. 28 Prototyp aplikace iDict (Hyrsykari a kol., 2000)

Na Katedře obecné lingvistiky Filozofické fakulty Univerzity Palackého v Olomouci se pracovníci věnují eye-tracking výzkumu zaměřenému na pacienty s agramatickou afázií a jejich vnímání pasivních vět. Velmi obsáhlou rešerši, ve které je komplexně shrnut eye-tracking výzkum v oblasti čtení, publikoval Rayner (1998).

### 2.2.8 Využití v marketingu a reklamě

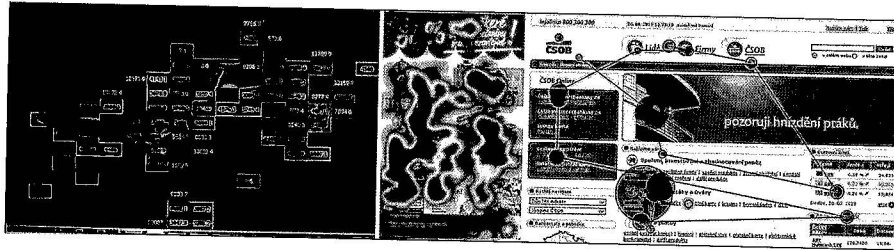
První známou studií využívající sledování pohybu očí v marketingu a reklamě je studie Nixona (1924), jenž pozoroval oči spotřebitelů, kteří si prohlíželi časopis obsahující tištěné reklamy, přičemž sám autor byl schovaný za závěsem. O několik let později využil Karlslake (1940) eye-tracking zařízení s názvem Purdue Eye Camera pro hodnocení reklam v novinách Saturday Evening Post. Klíčovým byl článek Russa (1978) s názvem „Eye-Fixations Can Save the World“, ve kterém nabádal k využívání eye-trackingu pro hodnocení efektivity marketingu a ke studiu rozhodovacích procesů spotřebitele.

Po publikování tohoto článku prudce narostl počet aplikací sledování pohybu očí v mnoha oblastech vizuálního marketingu. Na počátku devadesátých let dvacátého století bylo ročně publikováno jen asi deset článků o vizuálním vyhledávání v reklamě. O dekádu později to už bylo více než 200 článků (Wedel a Pieters, 2008).

Dnes je možné eye-tracking v marketingu rozdělit do několika oblastí. Využívá se pro analýzu chování spotřebitelů při výběru konkrétního zboží,

v tištěných i televizních reklamách, v politickém marketingu, designu etiket a brandingů a v současné době především v hodnocení webových stránek. Obsáhlou rešerší na toto téma vypracovali Wedel a Pieters (2008).

Hodnocení webových stránek je dnes nejčastějším polem, na němž jsou eye-trackery v komerční sféře využívány. V Česku se komerčním využitím eye-trackingu zabývá mnoho společností. Namátkou lze zmínit například UXFocus,<sup>19</sup> sledujemeoci.cz,<sup>20</sup> nebo Constat.<sup>21</sup> Z webových stránek těchto společností je patrné, že komerčně nabízené výstupy jsou ve většině případů omezené na obrázky attention map, Gridded AOI či průběhu scanpath (obrázek 29).



Obr. 29 Výstupy hodnocení stimulů z oblasti reklamy a marketingu<sup>22</sup>

### 2.2.9 Využití pro hodnocení použitelnosti

Výše popsané použití eye-trackingu v reklamě a marketingu lze považovat za součást hodnocení použitelnosti obecně. Nicméně marketingové využití eye-trackingu je natolik široké a časté, že bylo vyčleněno do samostatné kapitoly.

Historický přehled 21 studií využívajících eye-tracking pro hodnocení použitelnosti publikovali Jacob a Karn (2003). Autoři uvádějí, že první takovou studií v historii bylo už výše zmíněné testování armádních pilotů při přistávání (Fitts a kol., 1950). I další popisované studie byly zaměřeny na testování pilotů či řidičů, viz kapitola 2.2.5. Je tedy patrné, že eye-tracking výzkum je silně provázán, a striktní dělení tedy nelze vytvořit. Valná většina novějších studií popsaných v článku Jacoba a Karn (2003) je zaměřena na testování počítačových aplikací či webových stránek.

Právě testování webových stránek je v současnosti nejčastější oblastí hodnocení použitelnosti, kde je eye-tracking využíván. Velmi důležitou součástí webových stránek jsou různé formuláře. Jejich design je klíčový

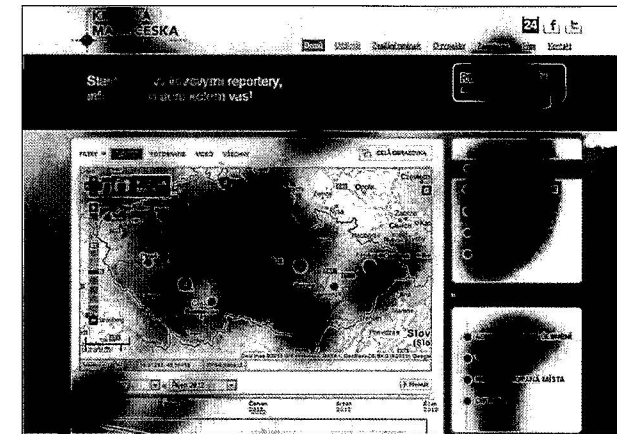
<sup>19</sup> <http://www.uxfocus.cz>

<sup>20</sup> <http://www.sledujemeoci.cz>

<sup>21</sup> <http://www.constat.cz>

<sup>22</sup> <http://www.constat.cz>

a eye-tracking může pomoci s odhalováním případných problémů. Jarret a Berstrom (2014) na základě několika případových studií uvádějí, že uživatelé prakticky nikdy nečtou instrukce, a proto je nezbytné vytvořit formulář co nejjednodušší tak, aby se v něm bez problémů zorientoval i laik. Zásadní je rovněž umístění tlačítek v designu stránky. To se prokázalo například ve studii Brychtové a kol. (2013), která hodnotila formulář pro zadání události do Krizové mapy Česka vytvořené Českou televizí. Cílem studie měla být evaluace vlastního formuláře. Jak se ale na základě naměřených eye-tracking dat ukázalo, respondenti měli největší problém s nalezením tlačítka pro zobrazení formuláře. Přestože toto tlačítko bylo velké a výrazné, respondenti si jej v layoutu stránky nevšimli. Bylo to způsobeno jeho umístěním v tmavém pruhu, který respondentům evokoval reklamu, a vůbec se do něj tedy nedívali. Attention mapa na obrázku 30 ukazuje rozložení fixací na analyzované stránce. Nejpomalejšímu z 15 testovaných respondentů trvalo 73 sekund, než si velkého žlutého tlačítka „Reportuj teď“ všiml.



Obr. 30 Ukázka nevhodně umístěného tlačítka pro reportování události v Krizové mapě Česka (Brychtová a kol., 2013)

Další velmi často testovanou oblastí jsou sociální sítě. Ty využívají miliony uživatelů a každé drobné vylepšení jejich použitelnosti má obrovský dosah. Burrige (2014) uvádí, že pro většinu sociálních sítí jsou klíčové obrázky. Jejich vhodným umístěním v layoutu stránky lze zvýšit její uživatelskou přívětivost. Burrige dále zmiňuje, že zajímavé by bylo porovnání trajektorií pohledu očí uživatelů sociálních sítí Twitter a Pinterest, protože Pinterest obsahuje především obrázky, zatímco Twitter je z velké části omezen na textové příspěvky.

S nárůstem počtu mobilních zařízení, jako jsou chytré telefony a tablety, vzrostl i zájem o hodnocení jejich použitelnosti. He a kol. (2014) uvádějí,



že nejprve se eye-tracking hodnocení mobilních zařízení omezovalo na simulaci interface na monitoru běžného zařízení. Výrobci ale poté eye-trackery upravili tak, že umožňovaly nahrávat data i z těchto zařízení. Posledním typem eye-trackerů využívaných v této oblasti výzkumu jsou mobilní brýle či headsety. Zajímavostí je, že i běžná kamera integrovaná v chytrém telefonu může být použita jako eye-tracker (Krafka a kol., 2016).

Velmi aktuální oblastí hodnocení použitelnosti pomocí eye-trackingu jsou počítačové hry. Kromě toho, že může být pohled použit k ovládní hry (viz kapitola 2.1), lze eye-tracking uplatnit i při hodnocení interface počítačových her či hledání rozdílů mezi zkušenými hráči a nováčky. Na Katedře geoinformatiky Univerzity Palackého v Olomouci proběhlo v rámci popularizace testování profesionálního hráče hry Counter Strike (obr. 31). Zajímavým zjištěním bylo, že hráč vůbec nesleduje zaměřovací kříž, ale soustředí se asi 3 cm nad něj. Z následného rozhovoru vyplynulo, že zkušený hráč přesně ví, kde je střed obrazovky, a tedy i zaměřovací kříž, a stačí mu vidět jej periferně.



Obr. 31 Testování profesionálního hráče hry Counter Strike

Oblastí hodnocení použitelnosti, kde lze eye-tracking využít, je mnohem více. Může se jednat například o hodnocení počítačových programů, diagramů, grafů, ale také map. A právě na hodnocení map jako součásti kognitivní kartografie je zaměřena následující kapitola.

## 3 KOGNITIVNÍ KARTOGRAFIE

### 3.1 Výhodiska kognitivní kartografie

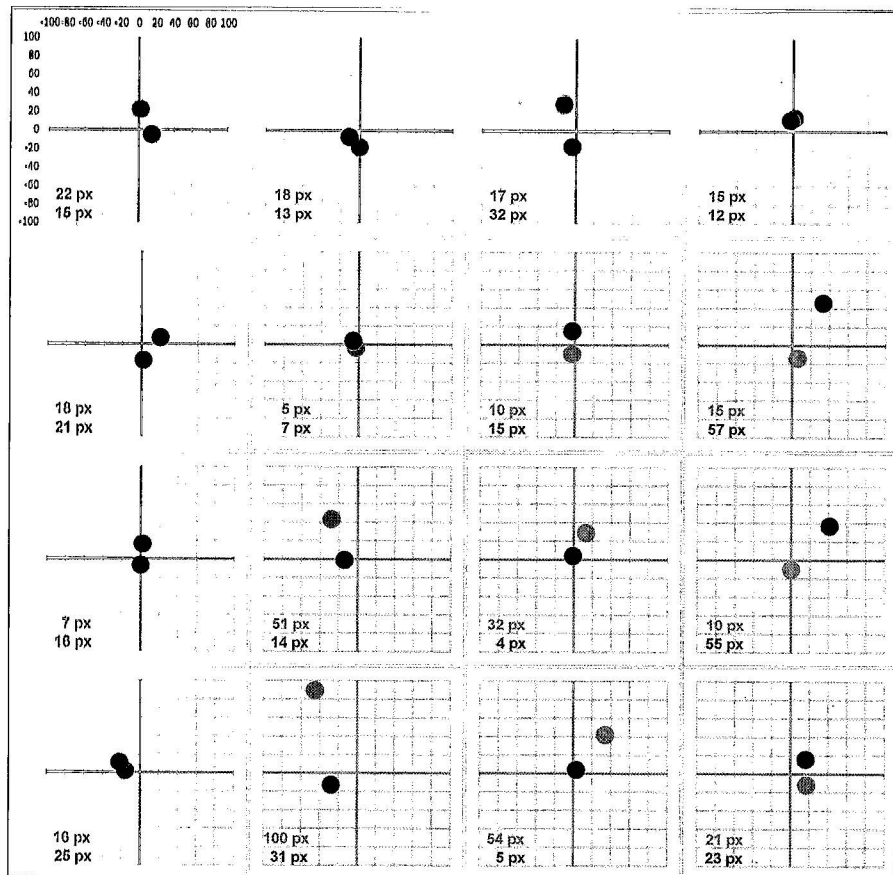
Díky masivnímu nárůstu produkce map během posledních let je stále důležitější zaměřit se na výzkum užití map. Na základě výsledků tohoto výzkumu lze mapy upravit tak, aby lépe sloužily uživatelům.

Elzaker (2004) rozlišuje dva základní typy testování v kartografii – funkční a kognitivní. Funkční testování užití map je založeno na předpokladu, že každá mapa je vytvořena pro určitý účel, a je proto důležité zjistit, zda a do jaké míry tento účel splňuje. V percepčním a kognitivním výzkumu užití map je častěji brán jako výchozí bod individuálního uživatele s cílem zjistit, „jak mapy fungují“. Kognitivní výzkum užití map tak odpovídá na otázku, „proč“ mapy a kartografické vyjadřovací prostředky pracují efektivně (Kubiček, 2011).

Důležité je rozlišovat mezi percepcí a kognicí. Percepci (vnímání) lze definovat jako organizaci a interpretaci sensorických informací. V kartografii percepce souvisí s primární reakcí uživatele na mapové vyjadřovací prostředky. Kognice však nezahrnuje pouze percepci, ale také proces myšlení, předchozí zkušenost a paměť. Dle slovníku cizích slov je kognice souhrn operací a pochodů, jejichž prostřednictvím si člověk uvědomuje a poznává svět i sebe samého. Výsledek těchto procesů je to, co člověk poznal prostřednictvím vnímání (percepce), učení, paměti, myšlení, uvažování, rozhodování, znalostí a komunikace (Montello, 2002). Těmito procesy probíhajícími při čtení mapy se zabývá kognitivní kartografie.

Kognitivní kartografii je možné členit do tří základních výzkumných směrů (Montello, 2002):

- *Map-design research* – jeho cílem je vylepšit mapy;
- *Map-psychology research* – jeho cílem je pochopit lidskou percepci a kognici;
- *Map-education research* – jeho cílem je vylepšit výuku s mapami a o mapách.



Obr. 68 Srovnání pozic fixací ve stimulu Slide2 pro čtrnáct účastníků. Vzdálenost od středu obrázku ukazuje fixační odchylku v pixelech. Data EyeTribe jsou zobrazena jako modré tečky, SMI data jako červené tečky.

## 5 PŘÍPRAVA EXPERIMENTU A ZÁZNAM DAT

Eye-tracking pro hodnocení použitelnosti produktů (map) lze využít dvěma způsoby. Jedná se o formativní (kvalitativní) a sumativní (kvantitativní) výzkum (Bojko, 2013).

Jak uvádí Hendl (2008), někteří metodologové chápou kvalitativní výzkum jako pouhý doplněk kvantitativních výzkumných strategií, jiní jako protipól nebo vyhraněnou výzkumnou pozici ve vztahu k jednotné, na přírodovědných základech postavené vědě. V typickém případě kvalitativní výzkumník vybírá na začátku téma a určí základní výzkumné otázky. Ty však může modifikovat nebo doplňovat v průběhu výzkumu, během sběru a analýzy dat. V průběhu kvalitativního výzkumu nevznikají pouze nové výzkumné otázky, ale také nové hypotézy a nové modifikace výzkumného plánu. Výzkumník vyhledává a analyzuje jakékoli informace, které přispívají k osvětlení výzkumných otázek, a provádí deduktivní a indukativní závěry. Často se stává, že své závěry probírá s účastníky výzkumu (respondenty) a jejich názory zohledňuje ve svých výsledcích. Pro pružný charakter kvalitativního výzkumu lze jeho výsledky jen těžko replikovat. Ve srovnání se statistickou analýzou se kvalitativní analýza nedá formalizovat (Hendl, 2008).

V kvalitativním (formativním) výzkumu je cílem výzkumníka nalezení oblastí, ve kterých by bylo možné daný produkt vylepšit. Příkladem může být hodnocení komplexní webové aplikace, kdy vývojář prezentuje hotovou verzi mapy skupině respondentů. Ti v aplikaci řeší zadané úkoly a na základě videozáznamu obrazovky a pohybu očí může vývojář zjistit, ve kterých úkolech respondenti chybovali, jestli se dívali na klíčové prvky aplikace, zda se vyznali v menu atd. Výzkumné otázky ve formativním výzkumu mohou znít například: „Všimnou si respondenti odkazu?“, „Budou se orientovat v menu?“ atd.

Druhým způsobem je sumativní neboli kvantitativní výzkum. Kvantitativní výzkumy v sociálních vědách v mnohém napodobují metodologii přírodních

věd. Kvantitativní výzkum používá náhodné výběry, experimenty a silně strukturovaný sběr dat. Získaná data jsou analyzována statistickými metodami s cílem je explarovat, popisovat, případně ověřovat pravdivost představ o vztahu mezi sledovanými proměnnými (Hendl, 2008).

Během kvantitativního (sumativního) výzkumu dochází k porovnání několika variant vytvořeného produktu. Příkladem může být opět hodnocení webové mapy. V tomto případě ale vývojář připraví dvě (nebo více) verzí mapy a následně testuje, která z nich je pro respondenty vhodnější, u které z nich vyřeší zadaný úkol rychleji, správněji atd. Výzkumné otázky v sumativním výzkumu mohou znít například: „Která varianta je vhodnější pro řešení úkolu?“, „Ve které variantě se respondenti rychleji orientují“ atd.

Základní vlastností kvantitativních experimentů je to, že výzkumník aktivně a úmyslně přivodí změnu určité situace (nebo hodnoceného produktu) a pak sleduje změnu chování respondentů. Vyjádřeno v jazyce kvantitativního výzkumu, výzkumník manipuluje s nezávislými proměnnými a měří změnu zvolených závislých proměnných. Tato problematika je popsána v následující kapitole.

## 5.1 Typy proměnných

Ke zjišťování kauzálních vztahů mezi proměnnými slouží experiment. Kauzální vztah znamená to, že změny v proměnné A jsou příčinami změn v proměnné B (Martin, 2007).

Aby bylo možné říct, že je mezi dvěma (nebo více) proměnnými kauzální vztah, musí být podle Campbella (1980) splněny přinejmenším tři základní podmínky:

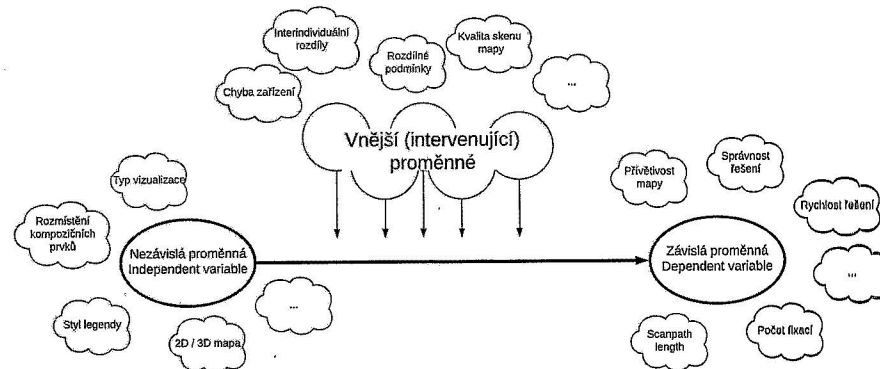
- 1) To, co je považováno za příčinu, musí časově předcházet předpokládanému efektu nebo následku.
- 2) Předpokládaná příčina i její efekt spolu musí kovariovat. Kovariance znamená společné změny. Jak se mění jeden znak (A), tak se zároveň v souladu s těmito změnami mění druhý znak (B).
- 3) Kromě vysvětlení změn jevu B proměnnou A nesmí existovat žádné jiné vysvětlení změn jinou proměnnou (např. C).

Splnění těchto tří podmínek zvyšuje hodnověrnost potenciálních závěrů o existenci nebo neexistenci kauzálních vztahů mezi proměnnými. Experiment, který tyto podmínky splňuje, bývá označován jako **vnitřně validní**. Zabezpečení maximální možné vnitřní validity je tedy základním požadavkem a závazkem pro každého výzkumníka (Punch, 2008).

Základními znaky vnitřně validního experimentu jsou tedy dle Ferjenčíka (2000):

- 1) manipulace s **nezávisle proměnnou** – předpokládanou příčinou;
- 2) měření **závisle proměnné** – předpokládaného efektu či důsledku;
- 3) kontrola všech jiných (**vnějších**) proměnných, které by mohly alternativně vysvětlovat změny závislé proměnné.

V příkladu z oblasti kartografie může být nezávislou proměnnou, se kterou je manipulováno, například stínování v mapě. Cílem výzkumníka je tedy zjistit, zda bude mít stínování vliv na měření závislé proměnné. Tou může být například doba řešení úkolu, správnost odpovědi nebo počet fixací na mapě. Intervenující proměnnou v tomto příkladu může být kvalita skenu u jednoho z typů hodnocených map. Výsledná zjištění, tedy například to, že na stínovaných mapách bylo zaznamenáno více fixací, tak mohou být způsobena nižší kvalitou skenu, nikoli použitím stínování, což chtěl výzkumník zjistit.



Obr. 69 Typy proměnných při tvorbě experimentu

Manipulace s nezávislou proměnnou je většinou jednoduchá. V kartografii se ve většině případů bude jednat o použití několika variant mapy lišící se v jedné – hodnocené – proměnné. Tedy například v tom, zda je legenda strukturovaná, či nikoliv, v použití různých barevných stupnic atd.

Měření závisle proměnné je také jednoduché. Měřit lze například rychlost odpovědi, správnost řešení, případně širokou škálu eye-tracking metrik, jako je například počet fixací, průměrná délka fixace či délka scanpath.

Největším a rozhodujícím problémem je kontrola nežádoucích (vnějších/ intervenujících proměnných). Ferjenčík (2000) ve své publikaci uvádí typologii vnějších proměnných.

Jako první typ vnější proměnné uvádí historii, tedy události a skutečnosti, kterými jsou respondenti ovlivněni paralelně s působením nezávisle proměnné. Ferjenčík popisuje, že pro eliminaci této proměnné je vhodné měřit závislou proměnnou bezprostředně po působení nezávisle proměnné. V oblasti eye-trackingu by s tímto neměl být problém, protože závisle proměnné (počet fixací atd.) jsou ze své podstaty měřeny ihned při pozorování stimulů.

Ani další typ vnější proměnné se eye-tracking výzkumu v kartografii příliš netýká. Touto proměnnou je zrání a přirozený vývoj. Pokud by byl například sledován vliv určité metody výuky po dobu několika let, je potřeba počítat s tím, že kromě této výuky na respondenty během oněch několika let působily i další vlivy.

Platí to i pro další typ vnější proměnné – reaktivitu pokusných osob – se zřejmě kartografických eye-tracking experimentů rovněž netýká. Efekt, kdy změna v závisle proměnné nastane spíše v důsledku očekávání než díky působení nezávisle proměnné, se nazývá placebo efekt.

Podobným typem vnější proměnné je očekávání ze strany experimentátora, kdy může experimentátor jakkoli nechtěně a nevědomky navádět subjekt k určitému chování. To může být problémem i u eye-tracking testování. K eliminaci této proměnné je vhodné, aby respondent na experimentátora neviděl, případně aby experimentátor vůbec nebyl ve stejné místnosti, což je však technicky komplikované.

Úspěch eye-tracking experimentu může zhatit i efekt měření, který se projeví u opakovaného testování stejného respondenta. Respondent může při druhém testování dosahovat lepších výsledků, avšak nikoliv proto, že jsou lepší prezentované stimuly, ale proto, že už ví, jakým způsobem bude test probíhat.

S opakovaným měřením souvisí i další z vnějších proměnných, a to experimentální mortalita. Takto bývá označován nechtěný úbytek respondentů v průběhu experimentu. Respondenti mohli testování vzdát například proto, že jim experiment přišel příliš složitý. Ve výsledku tak budou v experimentu pouze ti, pro které byl test jednoduchý.

Dalším, a nyní opravdu velmi významným, typem vnější proměnné je neekvivalenčnost skupin. Pokud testování probíhá tak, že jsou využívány dvě skupiny respondentů (between-subject design), je třeba zajistit ekvivalenci obou skupin, což není jednoduché.

Tento výčet intervenujících proměnných rozhodně není kompletní. Intervenující proměnnou je cokoli, kromě nezávisle proměnné, co může respondenta ovlivnit, a mít tak vliv na měření závisle proměnné.

Existuje několik přístupů, pomocí nichž je možné se vyhnout problémům s intervenujícími proměnnými. Nejjednodušším přístupem je prostá eliminace dané proměnné. Pokud by například byli respondenti během experimentu vyrušováni zvukem či světlem, je nejjednodušším řešením zatemnění oken

a odhlučnění místnosti, ve které experiment probíhá. Ferjenčík (2000) jako vzor uvádí I. P. Pavlova, jenž své pokusy dělal v laboratoři zvané věž mlčení, která byla téměř dokonale izolována od vnějšího světa. Tuto jednoduchou metodu bohužel není možné použít vždy. Těžko jde například eliminovat proměnné, jako jsou zkušenosti či motivace.

Druhým přístupem je **stabilizace a zkonstantnění**. Když není možné vnější proměnnou vyloučit, lze její hodnotu alespoň udržet na stejné úrovni během celého experimentu. Příkladem může být nastavení konstantního osvětlení v laboratoři.

Ve skutečnosti působí současně desítky až stovky různých faktorů (proměnných) a je technicky velmi obtížné je eliminovat či zkonstantnit. V případě, že by se to povedlo, však dochází k dalšímu problému, a tím je nedostatečná **externí validita**. Měřítkem externí validity je možnost a spolehlivost zevšeobecnění závěrů (Martin, 2007). V případě, že bychom eliminovali všechny intervenující proměnné, byly by výsledky experimentů příliš „umělé“ a jejich závěry by nebylo možné aplikovat do reálných situací.

Jak bylo zmíněno výše, jednou z nejproblematictějších vnějších proměnných jsou interindividuální rozdíly. Pokud jsou v experimentu porovnávány dvě skupiny respondentů, může se stát, že naměřené rozdíly mezi oběma skupinami nebudou způsobené nezávisle proměnnou, ale budou výsledkem toho, že členové jedné skupiny se ve svých charakteristikách významně liší od členů skupiny druhé.

Pro kontrolu interindividuálních rozdílů lze využít metodu **vyrovnávání**. Jejím cílem je vytvoření skupin, které budou rovnocenné nebo ekvivalentní. Vyrovnání není možné dosáhnout u všech vnějších proměnných. Proto je důležité zaměřit se na určité vytipované proměnné, které by mohly potenciálně kontaminovat vztah mezi závisle a nezávisle proměnnou. Přesto se však může stát, že výzkumník na nějakou klíčovou charakteristiku zapomene, nevyrovná ji, a ona pak ovlivní výsledky experimentu. Zároveň je velmi obtížné nalézt například dvacet respondentů, kteří se budou shodovat ve třech, čtyřech, či dokonce deseti vlastnostech (Ferjenčík, 2000).

Díky těmto nedostatkům se málokdy metoda vyrovnávání používá samotně. Mnohem častěji se aplikuje metoda **znáhodňování**. Její podstata vychází z teorie pravděpodobnosti. Pokud se podaří zajistit, že každý subjekt bude mít stejnou pravděpodobnost, že se dostane do kterékoliv pokusné skupiny, bude dosaženo maximální pravděpodobnosti, že všechny zkoumané skupiny budou z hlediska všech možných vlastností subjektů rovnocenné. Je důležité mít na paměti, že znáhodňování nezabezpečuje rovnocennost skupin. Zvláště při malém počtu pokusných osob ve skupinách se může stát, že tyto skupiny ekvivalentní nebudou.

Poslední metodou odstranění problémů s vnějšími proměnnými je jejich **transformace na nezávisle proměnné**. Tím pádem dojde k vytvoření multivariantského experimentu, protože ten bude obsahovat více nezávisle

proměnných. Takovýto experiment bude reálnější a bude mít vyšší externí validitu. Bude však mnohem náročnější na organizaci, přípravu výzkumného plánu a množství pokusných osob. Při rozšíření o druhou nezávisle proměnnou bude například nutné vytvořit čtyři skupiny respondentů.

## 5.2 Design experimentu

Jak bylo uvedeno v předchozí kapitole, podstatu experimentu tvoří tři podmínky: manipulace s nezávisle proměnnou, měření závisle proměnné a kontrola vnějších proměnných. Tím, jak splnit všechny tři podmínky, se zabývají plány experimentální plány či design experimentu. Cílem je vytvořit externě i interně validní a zároveň co nejefektivnější experiment. Efektivitou se rozumí úspora času, počtu respondentů a s tím související pracnosti.

Dle Ferjenčíka (2000) musí plánování experimentu začínat otázkou „co“:

Co bude **nezávisle proměnnou** – proměnnou, s níž bude manipulováno a o které výzkumník předpokládá, že bude příčinou nějakého jevu. V kartografii to může být například ověření nutnosti strukturování legendy a následná tvorba dvou (či více) variant mapy lišící se právě v použité legendě (např. strukturované a nestrukturované).

Co bude **závisle proměnnou** – proměnnou, u které se předpokládá, že se bude měnit v důsledku změn nezávisle proměnné. Ve výše uvedeném příkladu to může být rychlost nalezení určitého znaku v legendě či délka trajektorie pohybu oka zaznamenaná při tomto hledání.

Co bude **vnějšími proměnnými** – tedy, co ještě může vysvětlit změny v závisle proměnné. Zde se může jednat například o interindividuální rozdíly či rozdíly v zobrazovaných mapách. Pokud by například strukturovaná legenda v mapě byla psána jiným fontem či velikostí písma, nelze rozhodnout, zda na rychlost nalezení mapového znaku mělo vliv strukturování legendy či jiné písmo.

Bez ujasnění těchto tří „co“ nelze v tvorbě experimentu pokračovat. Klíčovou roli při tvorbě designu experimentu hraje počet nezávisle proměnných, se kterými bude manipulováno. U **jednoduchých plánů** experimentu dochází k manipulaci pouze jedné nezávisle proměnné. Opakem jsou tzv. **multivariační plány experimentů** (též multifaktorové či komplexní).

Jednoduché plány experimentů jsou, jak už název napovídá, jednodušší, transparentnější a zpravidla méně náročné na organizaci a vztah mezi závislou a nezávislou proměnnou je možné snáze interpretovat. Tyto plány jsou však vzdálenější realitě, kde souvisí všechno se vším (Ferjenčík, 2000). Ve své nejprostší variantě pracuje jednoduchý plán experimentu pouze se dvěma úrovněmi nezávisle proměnné (strukturovaná/nestrukturovaná legenda). Ve složitějších verzích může nezávisle proměnná nabývat více úrovní (stupnice s 5, 6, 7, 8 kategoriemi).

Mnohem důležitější rozhodnutí, než kolik úrovní bude mít nezávislá proměnná, je ale rozhodnutí o tom, jak přiřadit respondenty k různým úrovním nezávisle proměnné.

V případě, že je každý subjekt podroben působení pouze jedné úrovně nezávisle proměnné, se jedná o mezisubjektový plán experimentu (**between-subject experiment**). Tam, kde každý subjekt „přijme“ (v oblasti eye-trackingu tedy „vidí“) všechny úrovně nezávisle proměnné, se jedná o vnitrosubjektový plán experimentu (**within-subject experiment**).

### 5.2.1 Between-subject experiment

Na následujícím příkladu je popsán jednoduchý kartografický between-subject experiment zkoumající vliv strukturování legendy na délku trajektorie pohybu oka. Nezávisle proměnná tedy odpovídá typu legendy a bude dále označována jako  $X$ . Jelikož má dvě úrovně, bude varianta bez strukturování označena jako  $X_1$  a strukturovaná varianta jako  $X_2$ . Měření délky trajektorie oka je označeno jako  $Y$ . Plán experimentu tedy bude vypadat následovně:

první skupina respondentů	$X_1$	$Y$
druhá skupina respondentů	$X_2$	$Y$

Vliv strukturování legendy na délku scanpath bude zjištěn porovnáním délky scanpath u první a druhé skupiny respondentů. První dva znaky experimentu tedy byly naplněny. Byla změněna úroveň nezávisle proměnné a byla změřena úroveň závisle proměnné.

Zbývá tedy třetí podmínka – a to kontrola nežádoucích proměnných. Kromě typu použité legendy mohou mít na délku scanpath vliv i další faktory. Nejvýznamnějším z nich jsou interindividuální rozdíly mezi subjekty. Samozřejmě se mohou vyskytnout i další problémy, například že první skupina byla testovaná ráno a druhá odpoledne, že se změnil experimentátor (případně jeho nálada) a mnohé další. Většinu těchto faktorů je možné odstranit standardizací podmínek. Zbývá tedy zajistit, aby si obě skupiny byly rovnocenné. Zlatým pravidlem při sestavování between-subject experimentu je **přiřazování respondentů do skupin na základě náhodného výběru**. Metod, jak náhodně přiřadit respondenty, je mnoho – hod mincí, tabulka náhodných čísel, hod kostkou atd. (Ferjenčík, 2000).

Znáhodňování (**randomizace**) bývá ve schématech označováno jako  $R$ . Schéma popisovaného experimentu tedy vypadá následovně:

$R$	$X_1$	$Y$
	$X_2$	$Y$

Takovýto plán experimentu již splňuje všechna tři základní pravidla tvorby experimentu. Znáhodněním  $R$  totiž kontroluje explicitně nejvýznamnější vnější proměnnou, tedy interindividuální rozdíly. Popsané schéma

však není jediné možné, Ferjenčík (2000) na podobném příkladu uvádí, že lze znáhodnění zkombinovat s **vyrovnáním (V)**. Obě skupiny mohou být vyrovnány na základě nějaké důležité vnější proměnné – například úrovně kartografických znalostí. Na základě kartografického testu se vytvoří páry lidí s přibližně stejným počtem bodů, čímž dojde k vyrovnání obou skupin vzhledem k jejich kartografickým znalostem. Následně bude zavedeno znáhodnění. V každé dvojici bude losem rozhodnuto, do které skupiny budou respondenti přiřazeni. Schéma takového experimentu vypadá následovně:

$$V R \frac{X_1}{X_2} \frac{Y}{Y}$$

V některých případech je výhodné provést před vlastním testováním takzvaný pretest – tedy měření závisle proměnné před experimentálním působením. V popisovaném příkladu by se tedy před vlastní tvorbou experimentu respondentům ukázal stimulus obsahující legendu mapy (jinou, než bude použita v experimentu). V této legendě by respondenti měli za úkol nalézt konkrétní symbol. Na základě délky trajektorie pohledu oka by došlo k vyrovnání skupin a až po tomto vyrovnání by byli respondenti náhodně rozděleni do dvou skupin. U tohoto postupu je třeba dávat pozor na to, aby pretest nezpůsobil efekt učení, a neovlivnil tak respondenty při vlastním experimentu. U eye-tracking testování by navíc bylo nutné, aby respondenti přišli na testování dvakrát – poprvé na pretest a pak na vlastní experiment, což z časových a organizačních důvodů není vhodné.

Pro statistické zpracování naměřených výsledků je vhodné, aby skupiny obsahovaly stejný počet respondentů. K tomu se může použít takzvané blokové znáhodňování (**bloková randomizace**). Při této technice nejsou respondenti vybíráni do skupin po jednom, ale po skupinkách odpovídajících počtu úrovní nezávisle proměnné. Pokud jsou tedy v experimentu pouze dvě úrovně (strukturovaná a nestrukturovaná legenda), byli by náhodně vylosováni vždy dva respondenti a u nich se pak losem rozhodlo, do které skupiny budou přiřazeni.

Výhodou při between-subject experimentu je to, že každý respondent vidí pouze jednu úroveň nezávisle proměnné (jednu variantu mapy). Tím pádem není negativně ovlivněn učení při vícenásobném opakování experimentu. Dále nedochází k nežádoucímu prohlédnutí záměru experimentátora a rovněž se zmírňuje riziko únavy a nepozornosti respondenta.

Zásadní nevýhodou je však vliv interindividuálních rozdílů, který i přes znáhodňování a vyrovnávání nelze vždy odstranit. Zejména v případě malého počtu respondentů či velmi heterogenní skupiny není možné vytvořit rovnocenné skupiny. Z povahy eye-tracking experimentů plyne, že není technicky možné testovat velké počty subjektů. V takovýchto případech výzkumníci preferují vnitrosubjektový plán experimentu.

## 5.2.2 Within-subject experiment

Při vnitrosubjektovém plánu experimentu neboli within-subject experimentu, respondent přijímá (vidí) všechny úrovně působení nezávisle proměnné (tedy všechny varianty mapy). Takový design je mnohem ekonomičtější, protože je respondent porovnáván sám se sebou a především naprosto spolehlivě řeší problém neekvivalentnosti skupin.

I tento přístup má ale svá omezení a problémy. Mezi nejdůležitější z nich patří dle Ferencíka (2000) **transfer, učení a vliv pořadí prezentace stimulů**. K nejefektivnějším metodám kontroly takovýchto nežádoucích proměnných patří znáhodňování (randomizace) nebo vyvažování, označované také jako protivyvažování neboli **counterbalancing** (Bojko, 2013). Tím je myšleno uzpůsobení pořadí prezentace jednotlivých úrovní nezávisle proměnné (stimulů) tak, aby toto pořadí negativně neovlivňovalo vnitřní validitu experimentu. Within-subject experimenty se podle způsobu vyvažování dělí na kompletní a nekompletní.

V případě kompletního vnitrosubjektového plánu je kontrola pořadí a transferu nejdůležitější. V tomto případě obsahuje plán všechna možná pořadí prezentace stimulů, přičemž každý subjekt absolvuje všechny tyto kombinace („Subject by subject counterbalancing“). Tento přístup bývá označován jako ABBA vyvažování. V nejjednodušší variantě kompletního vnitrosubjektového experimentu (obsahujícího pouze dvě úrovně nezávisle proměnné) by byl respondentovi prezentován nejprve stimulus A, potom B a následně stejné stimuly v obráceném pořadí – tedy nejprve B a potom A. V příkladu se strukturovanou legendou by tedy byla respondentovi nejprve prezentována nestrukturovaná varianta (A), poté strukturovaná varianta (B) a následně ty stejné stimuly znovu, tentokrát ovšem v opačném pořadí (B a poté A). V případě, že by experiment obsahoval více úrovní nezávisle proměnné, stoupala by délka (a náročnost) experimentu geometrickou řadou.

Druhou variantou within-subject experimentu je nekompletní vnitrosubjektový plán. V tomto případě vidí každý respondent každou úroveň nezávisle proměnné pouze jednou. Aby se snížilo riziko transferu, učení nebo pořadí, jednotlivým respondentům jsou stimuly prezentovány v různých pořadích (Across-subject counterbalancing). Musí ale platit zásada, že každá podmínka (úroveň nezávisle proměnné) se musí na každém pořadovém místě objevit stejně často. K vyřešení této podmínky se nejčastěji používá tzv. latinský čtverec. Latinský čtverec je čtvercová tabulka o  $n \times n$  polích, která je vyplněna  $n$  různými symboly tak, že v každém řádku i v každém sloupci se každý symbol nachází právě jednou. V tomto případě budou těmito symboly úrovně nezávisle proměnné. Pokud by tedy bylo manipulováno se čtyřmi úrovněmi nezávisle proměnné (čtyřmi variantami mapy), mohl by latinský čtverec vypadat tak, jak je znázorněno na obrázku 70. V tomto konkrétním případě se jedná o vyvážený latinský čtverec, protože všechny úrovně nezávisle proměnné se vzájemně předcházejí a následují za sebou stejně často.

A	B	C	D
B	C	A	D
C	D	B	A
D	A	C	B

Obr. 70 Ukázka vyváženého latinského čtverce

Na obrázku 71 je znázorněn počet kombinací při vyvažování pro různý počet úrovně nezávisle proměnné při within-subject designu experimentu. Při kompletním vnitrosjektovém plánu je potřeba počítat s tím, že každý respondent uvidí všechny zobrazené kombinace. Pokud by předmětem testování byly například čtyři varianty legendy, musel by každý respondent absolvovat prezentaci 24×4 stimulů. Oproti tomu při vyrovnání pomocí latinského čtverce by byly každému respondentovi prezentovány pouze čtyři varianty (A, B, C a D). Každý respondent by je však viděl v jiném pořadí.

Úrovně nezávisle proměnné	Kompletní within-subject counterbalancing	Nekompletní within-subject counterbalancing
	Každý subjekt vidí VŠECHNY kombinace	Každý subjekt vidí pouze JEDNU kombinaci
2	AB BA	AB BA
3	ABC BCA CAB	ABC    CBA BCA    ACB CAB    BAC
4	ABCD    CABD ABDC    CADB ACBD    CBAD ACDB    CBDA ADBC    CDAB ADCB    CDBA BADC    DACB BACD    DABC BCDA    DBCA BCAD    DBAC BDCA    DCBA BDAC    DCAB	ABDC BCAD CDBA DACB
5	120 variant	ABECD    DCEBA BCADE    EDACB CDBEA    AEBCB DECAB    BACED AEDBC    CBDAE

Obr. 71 Kompletní a nekompletní counterbalancing u within-subject experimentů

Vyvažování a znáhodňování je v oblasti eye-trackingu technicky komplikované. V případě kompletního counterbalancingu by samozřejmě bylo možné vytvořit experiment prezentující stimuly v požadovaném pořadí. To by ale znamenalo, že v případě dvou úrovní působení nezávisle proměnné by byl stimulus prezentován celkem čtyřikrát, v případě čtyř úrovní už 24×4. To by mělo velký obrovský vliv na learning effect.

V případě znáhodňování pomocí latinského čtverce by bylo nutné pro každého respondenta vytvořit speciální, předem určené pořadí prezentace stimulů. To v softwarech dodávaných k eye-trackerům není možné, a bylo by tedy nutné použít aplikaci Hypothesis (viz kapitola 3.5.1) nebo nástroje, jako jsou komerční e-Prime<sup>52</sup> nebo open-source alternativa PsychoPy.<sup>53</sup> Tyto nástroje lze propojit s eye-trackerem a vytvářet v nich složitější a komplexnější experimenty. V případě, že chce výzkumník pro přípravu experimentu použít softwaru dodávaného k eye-trackeru, bude se muset spokojit s prostou randomizací prezentace snímků.

Jak bylo uvedeno výše, hlavním problémem between-subject designu experimentu byly interindividuální rozdíly mezi respondenty. Z toho důvodu musí u tohoto typu experimentu docházet k systematickému přiřazování respondentů do skupin, například pomocí náhodného výběru či pomocí vyrovnávání. U within-subject experimentu k tomuto problému nedochází, protože stejní respondenti vidí všechny varianty nezávisle proměnné. Je však třeba vyrovnat se s jiným zásadním problémem, a tím je efekt učení, označovaný také jako carryover nebo learning effect.

U některých typů experimentů může k odstranění learning effectu stačit výše popsány counterballancing nebo randomizace stimulů. Pokud je úkolem respondenta například hledání bodového znaku ve dvou variantách mapy, lze learning effect (alespoň částečně) odstranit prezentací stimulů v náhodném pořadí. Zejména pokud je stimulů vyšší počet.

Někdy je ale learning effect natolik zásadní, že toto samotné opatření nestačí. Příkladem může být nalezení státu s nejvyšší hodnotou zobrazovaného jevu. Pokud tento stát respondent nalezne na první variantě stimulu, jeho název si pravděpodobně zapamatuje a při zobrazení další varianty už bude odpověď znát, aniž by se musel dívat na mapu. V tomto případě je možné přistoupit k variantě, kdy zadání úkolu na obou (či více) variantách nebude totožné, ale bude co nejvíce podobné. Například ve studii Popelky (2014a) bylo cílem respondentů nalézt vrchol či obec s konkrétním názvem na stínované a nestínované mapě. Z důvodu eliminace learning effectu respondenti na obou variantách mapy nevyhledávali stejný vrchol, ale vrchol o podobné významnosti, popsány v mapě stejným fontem a stejnou velikostí písma. Kromě toho bylo dbáno na to, aby bod například neležel přímo ve středu snímku, na jeho okraji apod.

52 <https://www.scienceplus.com/e-prime-2-0-professional>

53 <http://www.psychopy.org/>

Jiný postup byl zvolen ve studii Popelky a Brychtové (2013), kdy respondenti řešili úkoly nad fiktivním terénem znázorněným pomocí barevné hypsometrie nebo pomocí perspektivní vizualizace. V tomto případě byla jedna mapa ve dvojici otočena o 180° nebo zrcadlově převrácena. Při následném rozhovoru s respondenty bylo zjištěno, že si během experimentu neuvědomili, že se jedná o stejné území.

Dalším řešením efektu učení může být realizace pauzy mezi dvěma testováními. Tento přístup byl využit například ve studii Opacha a kol. (2017), kde respondenti řešili úkoly nad dvěma variantami glyphů. Respondentovi byla náhodně přiřazena jedna varianta glyphů. Následovala pauza v délce alespoň tři dnů a poté respondent přišel na druhé testování, během něhož mu byla prezentována druhá varianta. Úkoly byly postavené tak, aby respondent během tří denní pauzy odpovědi zapomněl, což se opět potvrdilo v následném rozhovoru po testování.

V některých případech je ale learning effect tak výrazný, že within-subject design nelze vůbec použít. Bojko (2013) uvádí příklad s testováním dvou variant návodu na výměnu duše u kola. Pokud se respondent z jedné varianty naučí duši vyměnit, je prezentace druhé varianty zbytečná. V kartografii se lze s podobným problémem setkat například u hodnocení interaktivních prvků webové mapy. V tomto případě je tedy nezbytné zvolit between-subject design experimentu.

### 5.2.3 Volba designu experimentu

Volba mezi within-subject a between-subject designem experimentu záleží na výzkumníkovi. Při výběru je nutné brát ohled na specifika daného experimentu. Konkrétně na to, zda v daném experimentu převažují pozitivní či negativní vybraného plánu (obr. 72). Jak už bylo zmíněno v předchozích kapitolách, hlavním problémem u between-subject designu jsou interindividuální rozdíly. Naopak u within-subject experimentu je tím hlavním problémem learning effect. Pokud je tedy cílem experimentu například porovnání dvou variant interface webové mapy, je vhodnější zvolit between-subject, neboť poznatky z první varianty by si respondenti zapamatovali a použili je při řešení úkolu nad druhou variantou. Díky tomu, že u within-subject experimentu všichni respondenti vidí všechny stimuly, může být jejich počet nižší. Naopak délka testování bude delší, protože respondenti uvidí minimálně dvakrát tolik stimulů než v případě between-subject experimentu. Jedním z kritérií výběru plánu experimentu může být i fakt, že u within-subject experimentu mohou respondenti přímo porovnat obě (nebo více) testovaných variant například pomocí subjektivního dotazníku.

	Within-subject experiment	Between-subject experiment
Definice	Všichni respondenti vidí všechny varianty	Každý respondent vidí pouze jednu variantu
Hlavní problém	Learning effect	Interindividuální rozdíly
Počet respondentů	Nižší všichni jsou v jedné skupině	Vyšší potřeba více skupin po X respondentech
Délka experimentu	Delší	Kratší
Porovnání variant	Respondenti mohou porovnat varianty mezi sebou	Respondenti nemohou porovnat varianty mezi sebou

Obr. 72 Shrnutí vlastností within-subject a between-subject experimentu

### 5.2.4 Multivarianční experiment

Doposud popisované jednoduché plány experimentů situaci svým způsobem zjednodušovaly. Z množství proměnných byla totiž vždy vybrána pouze jedna, která byla považována za příčinu jevu. S touto nezávisle proměnnou bylo manipulováno a bylo zjišťováno, zda (a jak) se mění závisle proměnná. Všechny ostatní proměnné, které mohly zkoumaný jev také ovlivňovat, byly přitom eliminovány, zkonstantněny či zrovnováženy. Takovýto přístup je poměrně jasný, bezpečný a přímočarý. Vyloučení všech dalších proměnných ale vzdaluje experiment od reality.

Dobře sestavené a pečlivě realizované jednoduché plány experimentů zaručují poměrně vysokou vnitřní validitu, ale jejich slabinou bývá **ekologická validita**, tedy platnost výsledků v reálných podmínkách. Ekologická validita experimentu je možné dosáhnout například měřením přímo v terénu, kde bude ale experiment vystaven zvýšenému riziku působení mnoha vnějších proměnných, které není možné efektivně kontrolovat.

Druhou možností pro zvýšení ekologické validity experimentu je zvýšení počtu proměnných, se kterými bude manipulováno. Takovéto experimenty bývají nazývány komplexní nebo multivarianční. V komplexním experimentálním plánu jsou úrovně jednotlivých nezávisle proměnných kombinované faktoriálně. To znamená, že každá úroveň nezávisle proměnné je kombinována s každou úrovní ostatních nezávisle proměnných (Ferjenčík, 2000).

Na rozdíl od jednoduchých plánů experimentů dovedou multivarianční experimenty zkoumat interakce mezi jednotlivými nezávisle proměnnými. K přesnému matematicko-statistickému určení míry hlavních efektů a míry interakce slouží analýza rozptylu neboli ANOVA. Požadavky na technické



provedení a organizační zabezpečení multivariánčních experimentů jsou velmi vysoké a zvláště náročné na počet testovaných subjektů. Tento počet by bylo možné snížit pomocí volby nekompletního within-subject experimentu s využitím latinského čtverce. Díky nekompletnímu plánu by ale ztratily hlavní výhodu multivariánčních experimentů, totiž získání informace o interakci nezávisle proměnných. U multivariánčních experimentů lze rovněž kombinovat within-subject a between-subject design. Takovéto experimenty jsou potom označovány jako smíšené.

Bližší informace o designu experimentu, včetně příkladů multivariánčních plánů experimentu, lze nalézt například v publikaci Ferjenčika (2000) nebo Martina (2007).

### 5.3 Respondenti a jejich počet

Po definování hypotéz a sestavení plánu experimentu čeká výzkumníka další nelehký úkol, a tím je vymezení základního souboru. **Základním souborem** je množina všech prvků patřících do okruhu osob, které mají být v daném experimentu zkoumány (Pelikán, 1998).

Základní soubor musí být přesně vymezen. Pokud je cílem výzkumníka například testovat mapovou aplikaci pro studenty střední školy, je třeba si ujasnit, jaký typ středních škol bude uvažován, zda půjde o studenty všech ročníků nebo jen vybraných, ve kterém roce bude výzkum proveden, zda budou testováni studenti z celého světa, pouze z Česka, z vybraného města atd. Základní soubor může ale být velmi rozsáhlý. Bylo by velmi obtížné otestovat všechny středoškolské studenty. Z toho důvodu se ze základního souboru pořizuje výběr.

**Výběr** neboli **vzorek** by měl být co nejpřesnější zmenšenou kopií původního souboru (Pelikán, 1998). Co možná největší přiblížení vzorku základnímu souboru je důležité především proto, aby na základě výzkumu uskutečněného na výrazně menším souboru bylo možné dělat obecnější závěry platné pro celý základní soubor.

I při sebedokonalejším systému sestavení výběru nelze dosáhnout vnitřní identity vzorku se základním souborem. Pro co největší snížení těchto rozdílů musí být respektovány určité principy tvorby vzorku. Jak uvádí Pelikán (1998), v některých případech se nelze vyhnout tvorbě **záměrného výběru**. Tím je každý výběr, který výzkumník realizuje na základě svých zkušeností, poznatků a úsudků. V tomto případě výzkumník do vzorku vybírá ty osoby, které považuje za typické, a v podstatě se jedná o výběr subjektivní. V tomto případě se u seriózního výzkumu očekává, že ve výsledném článku budou použity formulace jako „... na daném vzorku se prokázalo...“ a výsledky nebudou vztahovány na celý základní soubor (Pelikán, 1998).

Zvláštní variantou záměrného výběru je **výzkum dobrovolně přihlášených osob**. Rosenthal (1965) uvádí, že se prokázalo, že dobrovolně

přihlášení respondenti mají vyšší intelektuální schopnosti, vyšší potřebu po sociálním uznání či nižší konvekčnost. Podobný výběr není možné využít například při standardizaci testů. Rovněž je vhodné ve výsledném článku na tento fakt upozornit.

Výrazně průkaznější tvrzení o obecné platnosti pro celý základní soubor lze učinit při zavedení **náhodného výběru** vzorku ze základního souboru. Základním rysem náhodného výběru je zaručení stejné pravděpodobnosti výběru pro každý prvek základního souboru. Kromě jednoduchého náhodného výběru lze využít i vícenásobný náhodný výběr. Ve výše uvedeném příkladu se studenty středních škol by byla nejprve náhodně vybrána střední škola, následně by byla náhodně vybrána jedna třída z dané školy a nakonec opět náhodně by z dané třídy byly vybráni respondenti.

Zásadní při přípravě experimentu je velikost vzorku. Tedy počet respondentů, kteří se experimentu zúčastní. Jednoznačné číslo, které by bylo univerzálně platné pro všechny typy experimentů, není možné stanovit. Před rekrutací respondentů je zásadní si nejprve uvědomit, zda se jedná o formativní (kvalitativní) či sumativní (kvantitativní) experiment.

#### 5.3.1 Počet respondentů ve formativním výzkumu

Jak už bylo uvedeno výše, cílem formativního výzkumu je zjistit problémy, na které respondenti narazí při práci s testovaným produktem. Sauro a Lewis (2016) uvádějí, že výpočet požadovaného vzorku pro formativní výzkum vychází z pravděpodobnostního modelu vyjádřeného rovnicí:

$$P(x \geq 1) = 1 - (1 - p)^n$$

pro pravděpodobnost nastoupení jevu **alespoň jednou při  $n$  náhodných pokusech** (například padnutí hlavy při hodu mincí), kde  $p$  vyjadřuje pravděpodobnost jevu. Nejedná se tedy o  $p$ -value při testování statistických hypotéz.

Pokud by například proběhlo pět hodů mincí, pravděpodobnost, že alespoň jednou padne hlava, by byla 0,969.

$$P(x \geq 1) = 1 - (1 - 0,5)^5 = 0,969$$

Tato rovnice může být převedena na rovnici pro výpočet velikosti vzorku. Protože  $n$  je v exponentu, je nutné rovnici logaritmovat:

$$\begin{aligned} (1 - p)^n &= 1 - P(x \geq 1) \\ n(\ln(1 - p)) &= \ln(1 - P(x \geq 1)) \\ n &= \frac{\ln(1 - P(x \geq 1))}{\ln(1 - p)} \end{aligned}$$

Aby bylo možné tuto rovnici použít, je nezbytné znát hodnoty  $p$  a  $P(x \geq 1)$ .

Nejpraktičtější postupem je zvolit nejnižší realisticky očekávané  $p$  s ohledem na možné zdroje. Sauro a Lewis (2016) uvádějí jako příklad formativní studii, kde je cílem dosáhnout alespoň 80% pravděpodobnost ( $P(x \geq 1)$ ), že bude minimálně jednou zpozorován problém, jehož pravděpodobnost pozorování je 0,15 ( $p$ ). Dle výpočtu je nutné pro takovou studii použít alespoň 10 respondentů.

$$n = \frac{\ln(1 - 0,8)}{\ln(1 - 0,15)}$$

$$n = \frac{\ln(0,2)}{\ln(0,85)}$$

$$n = 9,9$$

Na základě tohoto výpočtu lze tedy vytvořit tabulku (tab. 2), ze které je možné odečíst hodnoty nutného počtu respondentů.

Tab. 2 Velikost vzorku nutná pro formativní studii (upraveno dle Sauro a Lewis, 2016)

		$P(x \geq 1)$ pravděpodobnost alespoň jednoho pozorování problému					
		0,5	0,75	0,85	0,9	0,95	0,99
p pravděpodobnost pozorování	0,01	69	138	189	230	299	459
	0,05	14	28	37	45	59	90
	0,1	7	14	19	22	29	44
	0,15	5	9	12	15	19	29
	0,25	3	5	7	9	11	17
	0,5	1	2	3	4	5	7
	0,9	1	1	1	1	2	2

Lewis a Sauro (2016) rovněž uvádějí tabulku (tab. 3) zobrazující hodnoty pravděpodobnosti alespoň jednoho pozorování pro jednoho až dvacet respondentů a různé pravděpodobnosti pozorování.

Tab. 3 Hodnoty pravděpodobnosti alespoň jednoho pozorování pro různé počty respondentů a různé pravděpodobnosti pozorování

p	n = 1	n = 2	n = 3	n = 4	n = 5
0,01	0,01	0,02	0,03	0,04	0,05
0,05	0,05	0,1	0,14	0,19	0,23
0,1	0,1	0,19	0,27	0,34	0,41
0,15	0,15	0,28	0,39	0,48	0,56
0,25	0,25	0,44	0,58	0,68	0,76
0,5	0,5	0,75	0,88	0,94	0,97
0,9	0,9	0,99	1	1	1
p	n = 6	n = 7	n = 8	n = 9	n = 10
0,01	0,06	0,07	0,08	0,09	0,1
0,05	0,26	0,3	0,34	0,37	0,4
0,1	0,47	0,52	0,57	0,61	0,65
0,15	0,62	0,68	0,73	0,77	0,8
0,25	0,82	0,87	0,9	0,92	0,94
0,5	0,98	0,99	1	1	1
0,9	1	1	1	1	1
p	n = 11	n = 12	n = 13	n = 14	n = 15
0,01	0,1	0,11	0,12	0,13	0,14
0,05	0,43	0,46	0,49	0,51	0,54
0,1	0,69	0,72	0,75	0,77	0,79
0,15	0,83	0,86	0,88	0,9	0,91
0,25	0,96	0,97	0,98	0,98	0,99
0,5	1	1	1	1	1
0,9	1	1	1	1	1
p	n = 16	n = 17	n = 18	n = 19	n = 20
0,01	0,15	0,16	0,17	0,17	0,18
0,05	0,56	0,58	0,6	0,62	0,64
0,1	0,81	0,83	0,85	0,86	0,88
0,15	0,93	0,94	0,95	0,95	0,96
0,25	0,99	0,99	0,99	1	1
0,5	1	1	1	1	1
0,9	1	1	1	1	1

Dosažení hodnoty pravděpodobnosti pozorování může probíhat i na sofistikovanější bázi, například na základě předchozích realizovaných studií používajících stejnou metodu a obdobný objekt hodnocení. Sauro a Lewis (2016) na množství studií ukazují, že velikost  $p$  se běžně ve výzkumu pohybuje od 0,03 do 0,46.

Další možností je realizace pilotní studie, což představuje příklad hypotetické studie od Sauro a Lewis (2016), kde bylo pozorováno 10 problémů s testovaným produktem a které se zúčastnilo 10 respondentů. Písmeno X v následující tabulce 4 označuje situaci, kdy konkrétní respondent pozoroval konkrétní problém.

Tab. 4 Výsledky hypotetické studie dle Sauro a Lewis (2016)

respondent	problém										součet	proporce
	1	2	3	4	5	6	7	8	9	10		
P01	X	X		X		X	X	X			6	0,6
P02	X	X		X		X		X			5	0,5
P03	X	X		X	X	X					5	0,5
P04	X	X		X			X				4	0,4
P05	X	X	X	X		X			X		6	0,6
P06	X	X	X					X			4	0,4
P07	X	X	X		X						4	0,4
P08	X	X	X		X		X				5	0,5
P09	X		X		X		X		X		5	0,5
P10	X		X		X		X		X	X	6	0,6
součet	10	8	6	5	5	4	5	3	3	1	50	
proporce	1	0,8	0,6	0,5	0,5	0,4	0,5	0,3	0,3	0,1		0,5

Souhrnná pravděpodobnost pozorování problému ( $p$ ) v této studii byla 0,5. V případě, že by byla před vlastním experimentem realizována pilotní studie na čtyřech respondentech, jejím výstupem by byla tabulka 5.

Tab. 5 Výsledky prvních čtyř respondentů hypotetické studie dle Sauro a Lewis (2016)

respondent	problém								součet	proporce
	1	2	4	5	6	7	8			
P01	X	X	X		X	X	X		6	0,86
P02	X	X	X		X		X		5	0,71
P03	X	X	X	X	X				5	0,71
P04	X	X	X				X		4	0,57
součet	4	4	4	1	3	2	2		20	
proporce	1	1	1	0,25	0,75	0,5	0,5			0,71

Z této tabulky je možné vyčíst, že čtyři respondenti odhalili pouze sedm problémů testovaného produktu. Hodnota  $p_{est}$  pro tyto čtyři participanty je tedy 0,71. Pro odhad výsledné hodnoty  $p$  je nutné nejprve vypočítat tzv. deflated adjustment.

$$p_{def} = \left[ \left( p_{est} - \frac{1}{n} \right) \left( 1 - \frac{1}{n} \right) \right]$$

$$p_{def} = \left[ \left( 0,71 - \frac{1}{4} \right) \left( 1 - \frac{1}{4} \right) \right]$$

$$p_{def} = 0,345$$

Dalším krokem je výpočet tzv. Good-Turing adjustment na základě počtu objevených problémů a počtu problémů, které byly objeveny pouze jedním respondentem.

$$p_{GT} = \frac{p_{est}}{\left( 1 + \frac{N_{once}}{N} \right)}$$

$$p_{GT} = \frac{0,71}{\left( 1 + \frac{1}{7} \right)}$$

$$p_{GT} = 0,621$$

Posledním krokem je průměrování hodnot  $p_{def}$  a  $p_{GT}$ :

$$p_{adj} = \frac{0,345 + 0,621}{2}$$

$$p_{adj} = 0,48$$

Výsledek výpočtu upravené hodnoty  $p_{adj}$  je velmi podobný hodnotě  $p$ , pokud by byla vypočítána na podkladě dat všech deseti respondentů z tabulky 4.

Výpočtem lze rovněž odhadnout, kolik problémů dosud nebylo respondenty objeveno.

$$P_{\text{discovery so far}} = 1 - (1 - p_{\text{adj}})^n$$

$$P_{\text{discovery so far}} = 1 - (1 - 0,48)^4$$

$$P_{\text{discovery so far}} = 0,927$$

Tímto odhadem se zjistilo, že již bylo objeveno 92,7 % problémů testovaného produktu, přičemž u prvních čtyř účastníků bylo zjištěno osm problémů.

$$N_{\text{problems available for discovery}} = \frac{7}{0,927} = 7,6$$

Odhadovaný počet dostupných problémů je po zaokrouhlení osm. První čtyři respondenti odhalili sedm z nich, měl by tedy zbývat jeden problém, který dosud objeven nebyl. Z tabulky 4 je patrné, že v hypotetické studii bylo ve skutečnosti zjištěno deset problémů. Výsledek se tedy o dva problémy liší. Je nutné pamatovat na to, že tyto výpočty jsou pravděpodobnostní, nikoliv deterministické, a že je nutné počítat s určitou chybou či nepřesností.

Dalším nutným krokem je zvolení vhodné hodnoty  $P(x \geq 1)$ , tedy pravděpodobnost alespoň jednoho pozorování problému neboli Problem Discovery Goal. Sauro a Lewis uvádějí, že z historických důvodů je zvykem volit tuto pravděpodobnost v rozmezí mezi 80–85 %.

Velmi známou poučkou je tzv. „Magic Number 5“, tedy tvrzení, že k odhalení 85 % problémů stačí pět respondentů. Autorem tohoto tvrzení je Nielsen,<sup>54</sup> který vycházel ze studie Nielsena a Landauera (1993), kde se průměrná hodnota  $p$  pohybovala okolo 0,31. Při dosazení této hodnoty do rovnice  $P(x \geq 1) = 1 - (1 - p)^n$  je skutečně výsledkem číslo 0,8436. Nielsen však nikdy neřekl, že by formativní studie měly být prováděny pouze na pěti respondentech. Jeho tvrzení spíše směřovalo k tomu, že je vhodné provádět iterativní testování. Mezi jednotlivými iteracemi je vhodné provést změny testovaného produktu, ideálně odstranění již objevených problémů (Sauro a Lewis, 2016).

U studií, které počítají s iterativním testováním, lze tedy výsledný počet nezbytných respondentů podělit počtem iterací. Sauro a Lewis (2016) opět uvádějí příklad hypotetické studie, ve které je cílem výzkumníků objevit 90 % problémů s pravděpodobností pozorování  $p=0,15$ . Na základě dat v tabulce 2 je patrné, že by k tomu mělo stačit 15 respondentů. Výpočet lze provést rovněž prostřednictvím online kalkulátoru.<sup>55</sup> Výsledek tohoto výpočtu vypadá tak, jak je znázorněno na obrázku 73.

54 <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>

55 [https://measuringu.com/problem\\_discovery/](https://measuringu.com/problem_discovery/)

Input	Results
Discover <input type="text" value="90%"/> of all Problems.	You would need to test on average <b>14.17</b> users to discover 90% of UI problems given the occurrence of a problem is <b>15%</b> .
Problem Occurrence <input type="text" value="0.15"/> (values between 0 and 1)	
<input type="button" value="Calculate"/>	

Obr. 73 Výstup z online kalkulačky pro výpočet nutného počtu respondentů ve formativním výzkumu

V plánu této hypotetické studie se rovněž počítá s třemi iteracemi. Výzkumníci se tedy rozhodli použít nejprve tři respondenty, v druhé iteraci čtyři a v poslední zbývajících osm.

Při použití tří respondentů lze objevit 39 % problémů (tab. 3), což je velmi daleko od požadovaných 90 %. Nicméně již tato první iterace s pouhými třemi respondenty může odhalit 58 % problémů s  $p = 0,25$ , a 88 % problémů s  $p = 0,5$  a 100 % problémů s  $p = 0,9$ . Náprava těchto problémů může pomoci při další iteraci. Na konci druhé iterace, které se zúčastní čtyři respondenti, již bude celkový počet respondentů 7 (3 + 4), a lze tedy předpokládat, že bude objeveno 69 % problémů s  $p = 0,15$ . Po této iteraci může opět dojít k opravě chyb testovaného produktu. Po poslední iteraci bude tedy odhaleno všech 90 % problémů s  $p = 0,15$ .

Příkladem použití tohoto výpočtu může být studie zaměřená na hodnocení mapové aplikace, která proběhla na KGI UP v prosinci 2017. Během této studie bylo otestováno 15 respondentů. Následně bylo ověřeno, zda je tento počet dostatečný pomocí výše uvedených výpočtů.

V tabulce 6 je znázorněno, které ze sedmi zaznamenaných problémů byly pozorovány u některého z 15 respondentů. Celkem bylo zaznamenáno 28 problémů z celkového počtu 105 (7 problémů × 15 respondentů), což je 27 %. Hodnota pravděpodobnosti pozorování problému je tedy  $p = 0,27$ . Při zvolení pravděpodobnosti alespoň jednoho pozorování problému (Problem Discovery Goal)  $P(x \geq 1) = 0,95$  odpovídá nezbytný počet respondentů číslu 14. K tomuto číslu lze dojít dosazením do výše uvedených vzorců, případně použitím online kalkulátoru.<sup>56</sup> Při prostém dosazení hodnot  $p$  a  $P(x \geq 1)$  je výsledkem kalkulátoru počet 10 respondentů. Při zadání celé matice (tab. 6) je upravená hodnota  $p_{\text{adj}} = 0,2$  a hodnota nezbytného počtu respondentů je 14. Tato změna je způsobena tím, že dva problémy byly zaznamenány pouze u jednoho respondenta. Výstup kalkulátoru při zadání celé matice vypadá tak, jak je znázorněno na obrázku 74.

56 [https://measuringu.com/problem\\_discovery/](https://measuringu.com/problem_discovery/)

Tab. 6 Výsledky respondentů ve studii zaměřené na hodnocení mapového portálu

respondent	problém							součet	proporce
	1	2	3	4	5	6	7		
P01				X		X		2	0,29
P02	X				X			2	0,29
P03	X		X			X		3	0,43
P04					X			1	0,14
P05					X	X		2	0,29
P06						X		1	0,14
P07						X		1	0,14
P08		X		X		X		3	0,43
P09			X		X	X		3	0,43
P10	X				X	X		3	0,43
P11			X		X	X		3	0,43
P12				X			X	2	0,29
P13								0	0,00
P14						X		1	0,14
P15						X		1	0,14
součet	3	1	3	3	6	11	1	28	
proporce	0,20	0,07	0,20	0,20	0,40	0,73	0,07		0,27

**Input**

Discover  of all Problems.

Total participants

Problems Discovered:

**Results**

Given 28 total problems and 2 unique problem(s), the adjusted problem occurrence is 0.2. Which is the avg of Normalization: 0.18 and GT: 0.2

For the goal of discovering 95% of all problems available for discovery, the recommended sample size is 14 participants.

if a user encountered the problem enter 1 otherwise enter 0 or leave the field blank.

	Prob.	Prob.	Prob.	Prob.	Prob.	Prob.	Prob.	Count	P
	1	2	3	4	5	6	7		
User 1	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text" value="1"/>	<input type="text"/>	<input type="text" value="1"/>	<input type="text"/>	2	0.286
User 2	<input type="text" value="1"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text" value="1"/>	<input type="text"/>	<input type="text"/>	2	0.286

Obr. 74 Výstup online kalkulátoru při zadání celé matice pozorovaných problémů

Výsledkem tohoto výpočtu je tedy to, že pro odhalení alespoň 95 % problémů s danou pravděpodobností pozorování je nezbytné otestovat alespoň 14 respondentů. Jak již bylo uvedeno, studie se zúčastnilo celkem 15 respondentů. Tento počet je tedy dostatečný a výsledky studie mohou být považovány za reprezentativní.

### 5.3.2 Počet respondentů v sumativním výzkumu

Obecně lze říci, že v sumativním výzkumu je potřebný větší vzorek respondentů, což se projeví na ceně a zejména v případě eye-tracking testování i na čase nezbytném pro testování.

Tradiční odhad velikosti vzorku využívá odhad rozptylu měřené veličiny a posouzení, jak přesné musí měření být. Tato přesnost zahrnuje velikost kritického rozdílu a požadovanou úroveň statistické spolehlivosti (Walpole, 1976). Pokud jsou tyto informace dostupné, jedná se již jen o matematickou mechaniku.

Sumativní výzkum je možné rozdělit do několika oblastí. V první z nich je cílem výzkumu odhadnout určitou hodnotu (estimate value), například jak dlouho bude respondentům trvat řešení úkolu nad mapou. Druhou oblastí je porovnání naměřených hodnot s určitou, předem danou hodnotou (compare with a benchmark). Tedy zjištění, zda délka řešení odpovídá například 15 sekundám. Poslední a nejvýznamnější oblastí je porovnávání alternativ (compare with alternative). V tomto případě budou porovnávány různé varianty, například různé typy map, a to pomocí within-subject nebo between-subject experimentů. Rovněž je rozdíl, zda jsou výstupem experimentu binární nebo nebinární data. Jelikož eye-tracking metriky nejsou binární, bude tato kapitola zaměřena pouze na výpočet potřebného počtu respondentů z nebinárních dat.

Sauro a Lewis (2016) odvodili základní vztah pro výpočet potřebného počtu participantů při odhadu hodnoty (estimate value):

$$n = \frac{t^2 s^2}{d^2}$$

kde  $t$  je kritická hodnota úrovně statistické spolehlivosti,  $s^2$  je rozptyl dat a  $d$  je kritický rozdíl.

Odhad rozptylu  $s^2$  lze určit například na základě předchozích studií či pomocí pilotního experimentu.

Jako  $d$  je označen kritický rozdíl, jenž je cílem detekce, tedy nejmenší rozdíl mezi hodnotou naměřenou v experimentu a očekávanou hodnotou měřené veličiny. Neexistuje žádný matematický postup k určení příslušné hodnoty  $d$ . Určení této hodnoty je tedy záležitost úsudku, a to buď na základě zkušenosti výzkumníka, nebo pomocí přístupu „co kdyby“. Tento přístup spočívá v tom, že si výzkumník řekne, co by se stalo, kdyby byl rozdíl naměřený

a očekávané hodnoty 1 %. Takovýto rozdíl obvykle neovlivní výsledky studie. Pokud je tedy tento rozdíl nevýznamný, co by se stalo, kdyby byl rozdíl 5 %? Stejně lze postupovat až do doby, kdy si výzkumník řekne, že by zvolený kritický rozdíl negativně ovlivnil výsledky studie. Odhady kritického rozdílu mohou rovněž pocházet z předchozích studií používajících stejnou metodu a podobný typ stimulů. Pokud žádné takové studie nejsou k dispozici, lze provést pilotní studii na několika málo respondentech. Pokud ani pilotní studie není možná, lze dle Diamonda (2001) definovat kritický rozdíl odvozením od směrodatné odchylky dle vztahu:

$$e = \frac{d}{s}$$

$$d = e \cdot s$$

kde  $e$  je effect size (Cohen, 1977). Effect size (velikost efektu  $e$ ) se typicky používá jako standardní měřítko velikosti výsledku, vypočtené vydělením kritického rozdílu ( $d$ ) směrodatnou odchylkou ( $s$ ). Motivací pro zavedení velikosti efektu bylo definovat určitý ukazatel, který je na rozdíl od hladiny statistické významnosti nezávislý na velikosti vzorku (Minium, 1970).

Jako  $t$  je označena kritická hodnota požadované hladiny statistické významnosti. Určení této hodnoty je opět otázkou úsudku výzkumníka. Hladina statistické významnosti se obvykle stanovuje v rozmezí 80 % a 99 %, nejčastěji je však udávána hodnota 90 % nebo 95 %. Kritická hodnota je závislá na počtu stupňů volnosti a právě hladině statistické významnosti. K určení kritické hodnoty ( $t$ ) lze využít například online kalkulačtor.<sup>57</sup>

Použití  $t$  v tomto procesu však představuje komplikaci. Jak už bylo zmíněno, kritická hodnota  $t$  rozdělení ( $t$ ), na rozdíl od kvantilu normálního rozdělení ( $z$ -score), závisí na počtu stupňů volnosti ( $df$ ), který závisí na velikosti vzorku, což je veličina, která má být vypočítána. Pro jednovýběrový  $t$ -test je počet stupňů volnosti dán vztahem:

$$df = n - 1$$

Diamond (2001) nabízí řešení s využitím iterací. Iterace je možné představit na příkladu inspirovaném publikací Saura a Lewise (2016). Cílem hypotetické studie je zjistit, jak dlouho bude trvat řešení úlohy nad mapou. Výzkumník očekává, že řešení úlohy bude respondentům trvat 15 s, a rozhodne se stanovit kritický rozdíl na 10 % neboli 1,5 s. Po provedení pilotní studie na pěti respondentech bylo zjištěno, že řešení trvalo 12, 14, 12, 20 a 16 sekund. Rozptyl  $s^2$  těchto pěti hodnot je 11,2. Výzkumník se rovněž rozhodne pro  $\alpha = 0,05$  hladinu statistické významnosti. Výzkumník chce tedy shromáždit dostatek údajů, aby si mohl být na 95 % jist, že výsledný odhad délky řešení úkolu se bude lišit maximálně o 1,5 s od očekávané doby řešení (15 s). Nyní má výzkumník informace o rozptylu ( $s^2 = 11,2$ ), a kritickém rozdílu  $d = 1,5$  a potřebuje zjistit kritickou hodnotu úrovně statistické spolehlivosti ( $t$ ).

57 <http://stattrek.com/online-calculator/t-distribution.aspx>

Prvním krokem je využití kvantilu  $1 - \frac{\alpha}{2}$  normálního rozdělení ( $z$ -score) namísto  $t$ . Pro hladinu významnosti 5 % je to  $z = 1,96$ . Tento výpočet lze provést například pomocí online kalkulačtoru.<sup>58</sup> V tomto případě je v kalkulačtoru nutné vybrat možnost „Two-sided“. Následuje výpočet vztahu:

$$n = \frac{z^2 s^2}{d^2}$$

Výsledek této rovnice je 19,6, což se zaokrouhlí na 20 respondentů. V následujícím kroku bude odhad upraven tak, že  $z$ -score bude nahrazeno  $t$ -scorem pro  $n = 20$ . Pro tento odhad bude použito  $n - 1$  (tedy 19 stupňů volnosti  $df$ ). Hodnota  $t$  v další iteraci bude 2,093, což lze vypočítat například pomocí online kalkulačtoru.<sup>59</sup> Opět dojde k výpočtu  $n$  dle vztahu:

$$n = \frac{t^2 s^2}{d^2}$$

Jako  $t$  bude nyní dosazena hodnota 2,093. Výsledek výpočtu je 21,8, což po zaokrouhlení znamená  $n=22$ . Do výpočtu  $t$  bude nyní dosazeno 21 stupňů volnosti ( $df = 22 - 1$ ), a hodnota  $t$  tedy bude 2,08. Opětovným výpočtem bude dosaženo výsledku  $n = 21,5$ , což po zaokrouhlení vede ke stejnému počtu respondentů jako v předchozí iteraci ( $n = 22$ ). Jelikož dvě po sobě jdoucí iterace vedly ke stejnému výsledku, je hodnota 22 respondentů finálním výsledkem.

V případě, že data vykazují pozitivní šikmost, což je obvyklé například u času dokončení, který prakticky nemůže být v řádu zlomku sekundy, ale může být libovolně dlouhý, je vhodné použít pro výpočet velikosti vzorku geometrický průměr a přirozený logaritmus naměřených hodnot. Hodnota rozptylu bude vypočítána z přirozených logaritmů naměřených hodnot. Hodnota  $d$  musí být v tomto případě vyjádřena vztahem:

$$d_{ln} = \ln(\bar{x} + d) - \ln(\bar{x})$$

Čím mají data větší šikmost, tím větší bude rozdíl v potřebném počtu respondentů. Využitím přirozeného logaritmu je obvykle dosaženo nižšího počtu potřebných respondentů (Sauro a Lewis, 2016).

V případě, že nechce výzkumník pouze odhadovat určitou hodnotu, ale **porovnávat naměřená data s předem danou hodnotou (benchmarkem)**, je potřeba do vztahu zahrnout i chybu 2. druhu (viz kapitola 7.8). Výsledný vztah tedy bude vypadat následovně:

$$n = \frac{(t_\alpha + t_\beta)^2 s^2}{d^2}$$

Stejně jako v předchozím případě budou kritické hodnoty  $t$  nejprve nahrazeny kritickou hodnotou normálního rozdělení ( $z$ -score). Při těchto

58 <https://measuringu.com/zcalc/>

59 <http://stattrek.com/online-calculator/t-distribution.aspx>

výpočtech výzkumník nejprve použije hodnotu hladiny statistické významnosti ( $1 - \alpha$ ) (např. 0,95) pro oboustranný test (Two-sided) a následně určí hodnotu pro jednostranný test (One-sided) pro sílu testu neboli power ( $\beta$ ). K těmto výpočtům je opět možné využít online kalkulačtor.<sup>60</sup> Síla testu (power) je důležitá zejména tehdy, když nedojde k zamítnutí nulové hypotézy. Pokud je síla testu malá, je oprávněné zamítnutí nulové hypotézy málo pravděpodobné, a je tedy velice malá šance na přijetí platné alternativní hypotézy. Doporučovaná minimální hodnota pro sílu testu je 0,7 (chyba druhého druhu má tak pravděpodobnost maximálně 30 %).<sup>61</sup>

Třetí oblastí sumativního výzkumu je **porovnání alternativ**. Tato problematika je v oblasti eye-trackingu nejrozšířenější, protože ve valné většině případů dochází k porovnávání několika variant stimulů, například map. Jak bylo popsáno v kapitole 5.2, existují dva základní způsoby designu experimentu – within-subject a between-subject. Z podstaty rozdílného designu experimentu plyne i různý počet respondentů, kteří by se měli testování účastnit.

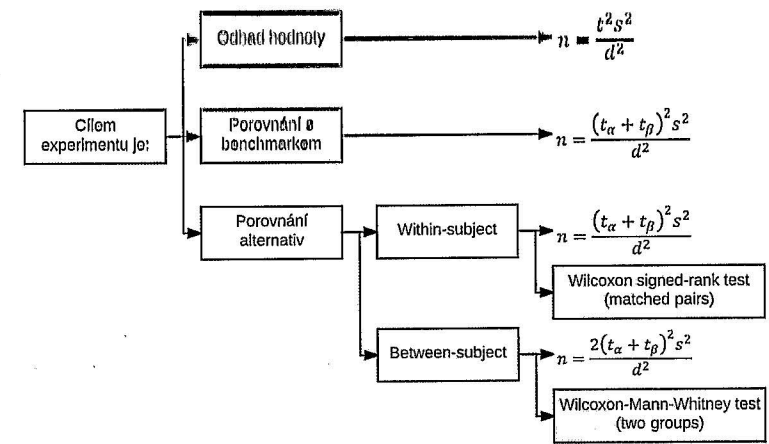
V případě within-subject experimentu vidí všichni respondenti všechny stimuly. V tomto případě je vztah pro výpočet počtu potřebných respondentů stejný jako v případě porovnávání s benchmarkem. Jedná se tedy o vztah:

$$n = \frac{(t_\alpha + t_\beta)^2 s^2}{d^2}$$

Between-subject experimentu se účastní (alespoň) dvě skupiny respondentů. Pro takový typ experimentu je tedy logicky nutné zajistit vyšší počet respondentů. Vztah pro výpočet tohoto počtu je opět velmi podobný předchozímu, pouze je doplněn o multiplikátor odpovídající počtu testovaných skupin. V nejčastějším případě se jedná o dvě skupiny (multiplikátor = 2), v případě, že by šlo o multivariacní experiment a testování by se zúčastnily například čtyři skupiny respondentů, měl by multiplikátor hodnotu 4. Rovněž je žádoucí, aby byly skupiny respondentů stejně velké. Výsledný vztah pro dvě skupiny respondentů je tedy:

$$n = \frac{2(t_\alpha + t_\beta)^2 s^2}{d^2}$$

Obrázek 75 shrnuje všechny důležité vztahy popsané v této kapitole. Je důležité pamatovat na to, že se jedná pouze o vztahy pro výpočet počtu respondentů u experimentů, jejichž výstupem jsou nebinární data.



Obr. 75 Přehled vztahů pro výpočet potřebného počtu respondentů

Dalším problémem je to, že všechny tyto vztahy jsou určeny pro data s normálním rozdělením. Hodnoty eye-tracking metrik naměřených během experimentu ale ve většině případů normální rozdělení nemají. Pro výpočet počtu potřebných respondentů u neparametrických testů je však možné použít například nástroj G\*Power<sup>62</sup> (Faul a kol., 2007). Modře zvýrazněné oblasti obrázku 75 popisují odpovídající volbu v programu G\*Power.

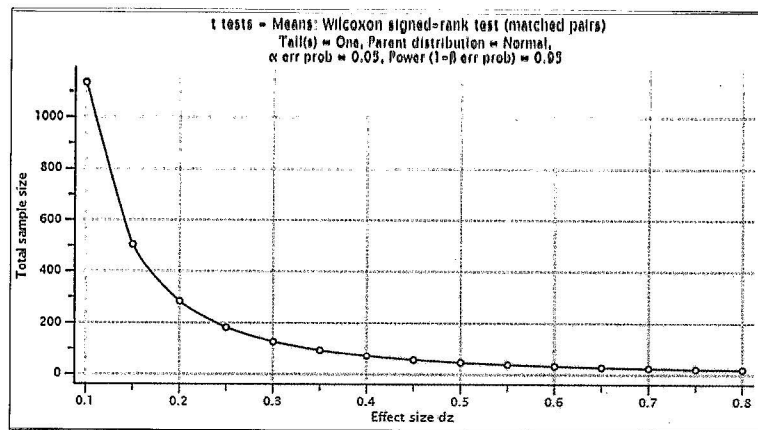
G\*Power 3 je nástroj pro výpočet síly mnoha statistických testů. Mezi ně patří F-test, t-test,  $\chi$ -test, z-test a další. Pro eye-tracking výzkum jsou důležité zejména dva testy z rodiny t-testů, a to Wilcoxon signed-rank test (matched pairs) pro within-subject experimenty a Wilcoxon-Mann-Whitney test (two groups) pro between-subject experimenty. Nástroj rovněž umožňuje několik typů analýz. Klíčové jsou opět dvě z nich, a to A priori a Post hoc.

U A priori analýzy je výsledkem potřebný počet respondentů, který je vypočítán z velikosti efektu, hladiny statistické významnosti a síly testu. A priori analýzu je tedy vhodné provádět před vlastním experimentem. Hladinu statistické významnosti (např.  $\alpha = 0,05$ ) i sílu testu (např.  $\beta = 0,7$ ) si výzkumník stanoví na základě svých zkušeností a požadavků. Klíčovou veličinou je tedy síla efektu. Cohen (1988) navrhl použití hodnot 0,2, 0,5 a 0,8 pro malé, střední a velké efekty. S rostoucí hodnotou síly efektu klesá počet požadovaných respondentů. Příkladem může být výpočet požadovaného počtu respondentů pro Wilcoxonův test při within-subject experimentu s nastavením pravděpodobnosti chyb 1. a 2. druhu dle obrázku 76.

60 <https://measuringu.com/zcalc/>

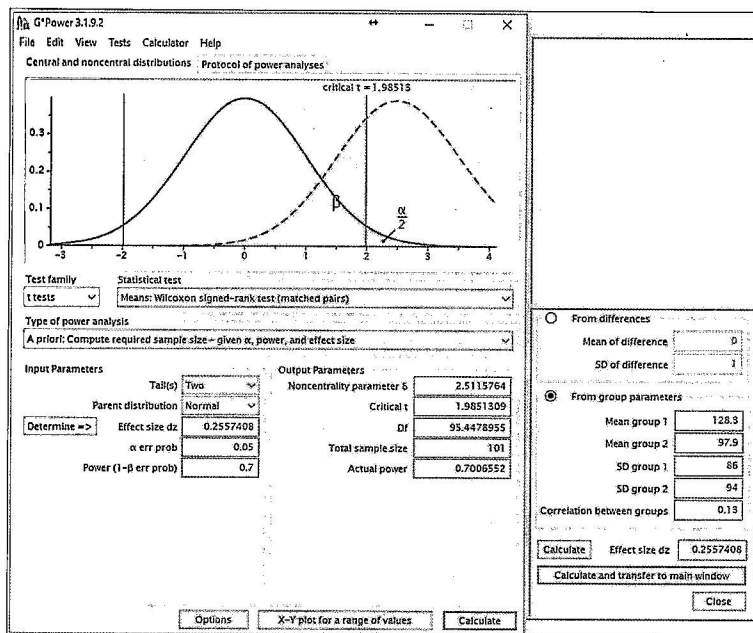
61 [http://www.statsoft.cz/file1/PDF/newsletter/2014\\_04\\_29\\_StatSoft\\_Sila\\_testu.pdf](http://www.statsoft.cz/file1/PDF/newsletter/2014_04_29_StatSoft_Sila_testu.pdf)

62 <http://www.gpower.hhu.de/>



Obr. 76 Výstup programu G\*Power pro Wilcoxonův test při within-subject experimentu

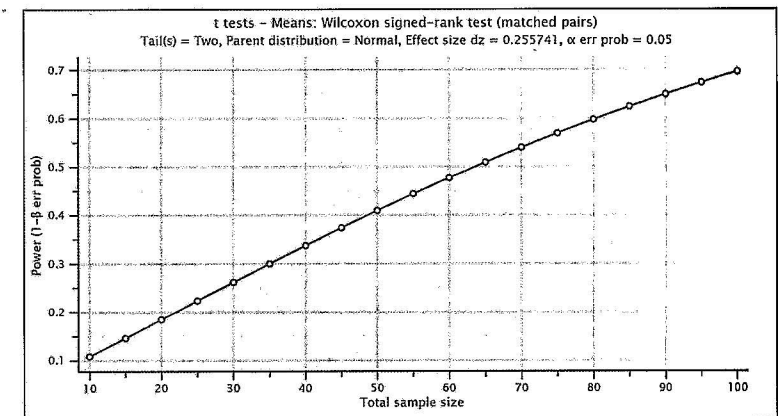
Při malé velikosti efektu (0,2) bude třeba otestovat 285 respondentů, při střední velikosti efektu (0,5) se jedná o 47 respondentů a při velké velikosti efektu (0,8) je třeba otestovat pouhých 20 respondentů.



Obr. 77 Výsledek A priori analýzy v programu G\*Power. Hladina statistické významnosti byla stanovena na  $\alpha = 0,05$ , síla testu na  $\beta = 0,7$  a velikost efektu byla odvozena od již proběhlé eye-tracking studie. Dle výsledku se mělo experimentu zúčastnit alespoň 101 respondentů.

Kromě prosté volby na základě Cohenova doporučení je možné velikost efektu odvodit například z výsledků podobné studie či pilotního experimentu. K tomu je nutné znát pro obě testované alternativy (například stínovaná a nestínovaná mapa) průměr a směrodatnou odchylku naměřených hodnot a také korelaci mezi oběma alternativami. Po zadání těchto informací do okna G\*Power se vypočítá velikost efektu a následně i potřebný počet respondentů. V případě vysokých hodnot směrodatných odchylek je hodnota velikosti efektu nízká a následně je třeba vysoký počet respondentů. Příkladem může být ukázka na obrázku 77, kdy byla velikost efektu vypočítána na základě již proběhlého eye-tracking experimentu, jehož se zúčastnilo 26 respondentů. Jak je z obrázku patrné, hodnoty směrodatných odchylek pro obě skupiny jsou vysoké, takže výsledná velikost efektu je pouze 0,26. Výsledkem A priori analýzy je tedy počet 101 potřebných respondentů, což představuje téměř čtyřnásobek skutečného počtu.

Druhým typem vhodné analýzy je Post hoc, tedy zpětný výpočet síly testu na základě použitého počtu respondentů, hladiny statistické významnosti a velikosti efektu. Po zadání všech hodnot (velikost efektu 0,257; hladina statistické významnosti  $\alpha=0,05$  a počet respondentů 26) dojde program k výsledku, že síla tohoto konkrétního testu byla 0,231. G\*Power umožňuje vykreslit graf, ve kterém je znázorněna síla testu v závislosti na počtu respondentů (obr. 78).



Obr. 78 Graf znázorňující sílu testu v závislosti na počtu respondentů

V případě between-subject experimentu by byl potřebný počet respondentů ještě vyšší, protože je třeba otestovat dvě skupiny respondentů. Ve většině experimentů je rovněž testováno více eye-tracking metrik. V tomto případě je vhodné vypočítat sílu testu pro všechny z nich a nejvyšší počet respondentů považovat za doporučený.



Velmi často je rovněž uváděno, že ideálním počtem respondentů je 30, případně 32. Autory tohoto doporučení jsou Nielsen a Pernice (2010), kteří testovali 60 respondentů nad vizuálním stimulem a následně testovali, při jakém počtu respondentů nebude rozdíl ve výsledné attention mapě. Při rozdělení do šesti skupin po deseti respondentech zjistili, že mezi šesti výslednými attention mapami jsou značné rozdíly. Obdobná situace nastala v případě dělení na tři skupiny po dvaceti respondentech. Až při rozdělení na dvě skupiny po třiceti respondentech se výsledné attention mapy jeví jako stejné. Autoři však attention mapy porovnávali pouze vizuálně a vůbec nehodnotili statistickou analýzu naměřených dat.

Jak uvádí Bojko (2013), výsledný počet respondentů ve skutečnosti vždy závisí na dostupných zdrojích. Zejména v případě eye-tracking testování je hlavním omezením čas, neboť ve většině případů může být najednou testován pouze jeden respondent. Z analýzy patnácti kvantitativních kartografických eye-tracking studií (Popelka, 2015) vyplynulo, že průměrný počet respondentů je 25. Nejvyšší počet 45 respondentů využil v kartografii Nevelsteen (2013). Přesto je vhodné před každou studií vypočítat potřebný počet respondentů a pokusit se tomuto počtu alespoň přiblížit. Stejně tak je vhodné ve finálním článku reportovat sílu testu, které bylo s daným počtem respondentů dosaženo.

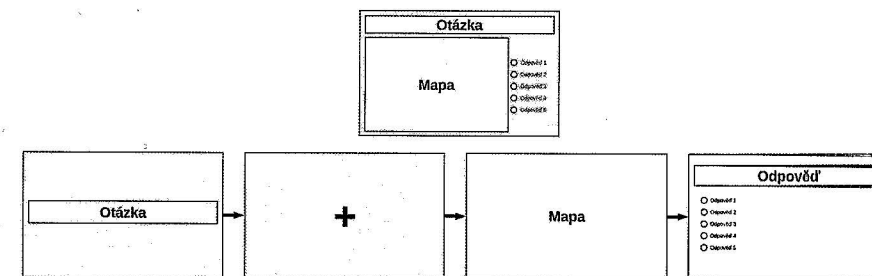
#### 5.4 Příprava stimulů

V prvních experimentech prováděných na KGI UP v roce 2011 byly otázky a možné odpovědi umísťovány přímo do stimulu s mapou. Výhodou tohoto přístupu bylo, že respondent si otázku a odpovědi nemusel pamatovat. Naměřená data ale byla velmi výrazně ovlivněna opakovaným čtením zadání úkolu, u kterého se respondent nesoustředil na vlastní obsah stimulu. Počet fixací byl tedy více ovlivněn množstvím textu v zadání než testovanou vizualizací. Z toho důvodu nebylo možné statisticky analyzovat eye-tracking metriky.

Vhodnější (pro kartografické experimenty) je tedy umístění otázky na samostatný stimulus prezentovaný před hodnocenou mapou. Samozřejmě že tento přístup není nejlepší ve všech experimentech bez výjimky. Například ve studiích zaměřených na hodnocení výukových materiálů je vhodné získat informaci, jak často se respondent očima vrací do zadání. Pokud je však předmětem testování mapa, je vhodné umístit otázku samostatně.

Otázky je vhodné koncipovat tak, aby byly co nejjednodušší a respondent si je zapamatoval. Ze zkušenosti se ukázalo, že respondenti mají problém si zapamatovat už dva pojmy. V jedné ze studií (Zieglerová, 2012) bylo testováno použití namluvené otázky, která byla respondentům přehrávána v průběhu prezentace stimulu s mapou. Tento přístup se však také neosvědčil. Přestože se respondenti na audiozáznam soustředili (což bylo možné rozpoznat z trajektorie pohybu jejich očí), mnohdy si otázku nezapamatovali.

Při přípravě eye-tracking experimentu je vhodné zajistit, aby trajektorie pohybu oka všech respondentů začínala vždy ve stejném místě stimulu, jinak jsou výsledky ovlivněny různým umístěním začátku trajektorie (každý respondent se začal na stimulus dívat z jiného místa). Jednoduchou metodou, jak toho docílit, je vložení samostatného stimulu s fixačním křížem uprostřed obrazovky před stimulus s úlohou (obr. 79). Tento stimulus postačuje zobrazit pouze na 500–600 ms a pohyby očí během jeho prezentace nahrávat. Respondenti se na fixační kříž intuitivně podívají a počátek trajektorie pohledu nastaví na střed stimulu.



Obr. 79 Nevhodně vytvořený stimulus (nahore) a doporučená varianta oddělující otázku a odpovědi od vlastního stimulu (dole)

Při umístění odpovědí přímo do stimulů nastává stejný problém jako u otázky. Pro analýzu dat je vhodnější využít dotazníku (např. stimulus typu question v SMI BeGaze) vloženého přímo do experimentu za vlastní stimulus, případně označení odpovědi pomocí kliknutí myši přímo do stimulu. Odpovědi respondentů lze následně vyexportovat ve formě tabulky a snáze je analyzovat.

#### 5.5 Tvorba experimentu

K většině eye-trackerů je dodáván software určený pro přípravu experimentu. Funkcionalita těchto aplikací je velmi podobná a pro většinu experimentů v oblasti kartografie dostatečná. V případě, že by chtěl výzkumník vytvářet sofistikovanější experiment, obsahující například větvení na základě zaznamenaných odpovědí, musel by sáhnout po specializovaném softwaru, jako je e-Prime<sup>63</sup> nebo volně dostupné PsychoPy.<sup>64</sup>

63 <https://www.scienceplus.com/e-prime-2-0-professional>

64 <http://www.psychopy.org/>

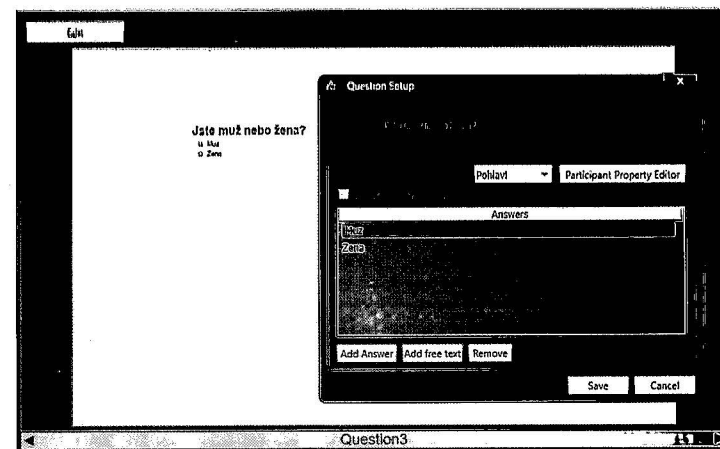
### 5.5.1 Tvorb experimentu v SMI Experiment Center

Software SMI Experiment Center umožňuje použít několik typů stimulů. Jedná se o text, PDF, obrázek, video, záznam obrazovky či web. Po dokončení modulu bylo možné využít i obraz externí kamery. Experiment lze rovněž doplnit o dotazník pomocí stimulu typu question.

Nejčastějším typem stimulu využívaným v kartografických experimentech je obrázek. Ten je vhodné připravit přímo v rozlišení, v jakém bude na stimulus monitoru prezentován. SMI Experiment Center samozřejmě umí zobrazit i obrázky v jiném rozlišení, ale mohlo by docházet k problémům při následném importu do aplikace OGAMA (viz kapitola 6.1). Výsledné stimuly (doplněné o pruhy po stranách tak, aby byly zobrazené na celý monitor) jsou sice uloženy ve složce Experiments, ale od verze 3.6 jsou pojmenovávány hash kódem, proto by bylo problematické jejich spárování s daty v OGAMA.

Dalším často využívaným typem stimulu je Screen Recording neboli záznam obrazovky. Záznam obrazovky je určen například pro analýzu práce s určitým programem, ale lze jej využívat i pro záznam práce s webovou stránkou. SMI Experiment Center sice obsahuje přímo stimulus typu web, ale jeho použití je limitováno na statické webové stránky. Pokud by součástí stránky bylo například okno s mapou, nebude pohyb této mapy zaznamenán. Obdobný problém nastává v případě dynamických menu. Při použití stimulu typu Screen Recording je nutné počítat s tím, že výsledkem je videozáznam pro každého respondenta zvlášť doplněný o trajektorii pohybu jeho očí. Analýza takovýchto dat může být velmi časově náročná, protože nelze srovnat pohyby očí více respondentů nad jedním stimulem. Každý z nich totiž s nahrávaným programem manipuloval jinak. V některých případech, například když je část stimulu statická a část dynamická, lze tuto nevýhodu obejít pomocí funkce Custom Trial Selector (viz kapitola 6.5).

Dalším často využívaným typem stimulu je Question neboli dotazník. Ten je vhodné použít pro získání osobních informací o respondentech, jako je jejich jméno, věk, vzdělání atd. Při použití tohoto dotazníku je velmi důležité nastavit propojení na pole v databázi, kam se bude vyplněná informace ukládat. Pokud by se tak nestalo, dotazník by sice fungoval, ale vyplněné informace by se nikam neuložily. Toto pole se vytváří v nabídce Participant Property Editor. Ta je dostupná buď přímo v okně, kde se vytváří dotazník, nebo přes Extras – Participant Property Editor. Dotazník v SMI Experiment Center umožňuje vybrat více než jednu odpověď. V takovémto případě se ale odpověď neuloží do databáze a výzkumník ji tedy nebude moci analyzovat. Proto je vhodné otázky volit tak, aby respondent označoval vždy pouze jednu možnost. Náhled dialogového okna je znázorněn na obrázku 80.



Obr. 80 Dialogové okno stimulu typu Question v SMI Experiment Center

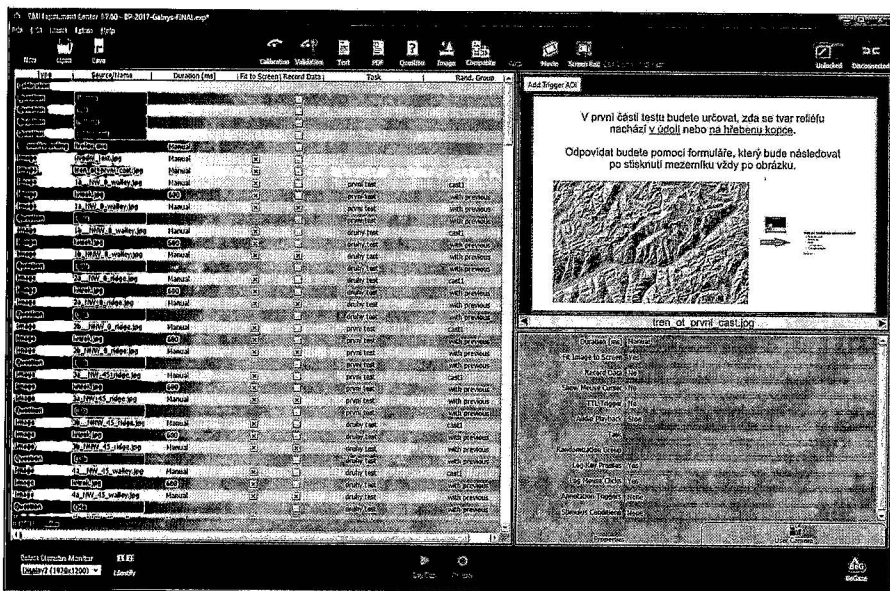
Posledním typem stimulu, který je běžně v kartografických experimentech využíván, je text. V kartografii se jeho použití omezuje téměř výhradně na formulaci otázek. Výzkumníci zabývající čtením jej využijí mnohem více. Aplikace SMI BeGaze totiž umožňuje zadaný text automaticky rozdělit na věty či slova, vytvořit kolem nich oblasti zájmu a následně analyzovat pohyby očí mezi těmito oblastmi.

Jak už bylo uvedeno výše, stimulus typu web je pro moderní internetové stránky téměř nepoužitelný. V kartografii rovněž nenajde využití stimulus typu PDF či video, které jsou však uplatnitelné v jiných disciplínách. Pomocí nabídky Composite lze kombinovat text, obrázek a video do jednoho stimulu, ale ani tato funkce v kartografii zřejmě využití nenajde.

Nezbytnou součástí každého eye-tracking experimentu je kalibrace. SMI Experiment Center nabízí volby od 0 do 13bodové kalibrace. Rovněž lze definovat rychlost pohybujícího se cíle a jeho styl. Zajímavá je možnost použití vlastního obrázku, což může být přínosné zejména u studií s malými dětmi. Kalibraci je vhodné doplnit o validaci (volba Quality Check) tak, aby bylo možné zkontrolovat, jaká je naměřená odchylka, a případně kalibraci zopakovat.

Úkázka typického eye-tracking experimentu je znázorněna na obrázku 81. V levé části je zobrazen seznam jednotlivých prvků (stimulů) experimentu, v pravé pak náhled vybraného stimulu a jeho vlastnosti. Prvním prvkem v experimentu je již zmíněná kalibrace následovaná čtyřmi osobními otázkami. Následuje jeden stimulus typu ScreenRecording. Ten je nastaven tak, že automaticky spustí internetový prohlížeč a v něm internetovou stránku obsahující vytvořenou databázi respondentů. Tento stimulus je na KGI UP vkládán do každého experimentu, takže je vytvářena databáze všech respondentů, včetně informace, kterých experimentů se zúčastnili.

Následují dva stimuly obsahující informace o experimentu a jeho otázkách. Vhodné je rovněž použití trénovačního stimulu, na kterém si respondent vyzkouší, jaké úkoly bude řešit. Dále následuje vlastní otázka a poté fixační kříž. Tento stimulus je nastaven tak, že bude zobrazen po dobu 600 ms. Doba trvání lze nastavit u všech stimulů, ale ve většině případů je vhodné použít volbu manual, kdy je stimulus zobrazen, dokud respondent nestiskne mezerník nebo některou z funkčních kláves, která je nastavena pro posun na další stimulus. Toto nastavení se provádí v dialogovém okně Extras – Global Settings. Po fixačním kříži je už zobrazen vlastní testovaný stimulus, v tomto případě obrázek. Jak je patrné z obrázku 81, až u tohoto stimulu jsou zaznamenávána eye-tracking data. Následuje stimulus typu Question, kde respondent vybírá odpověď z předem připravených možností. Další možností je odpověď pomocí kliknutí přímo do mapy. V tomto případě je nezbytné povolit zobrazení kurzoru a logování kliknutí myši.



Obr. 81 Typický kartografický eye-tracking experiment v prostředí SMI Experiment Center

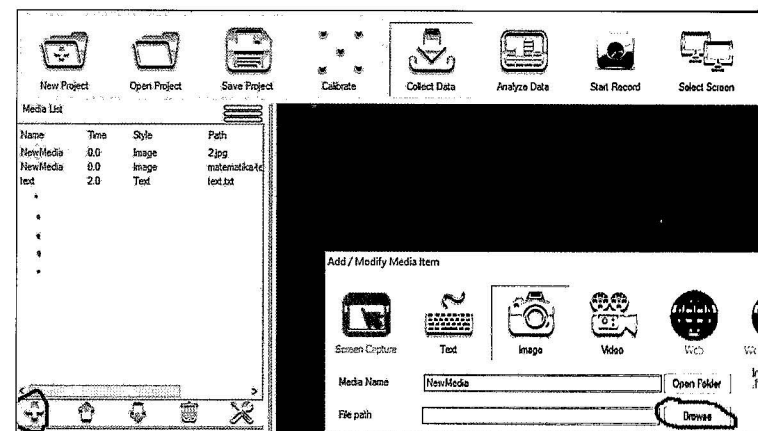
Poslední dva sloupce v přehledu experimentu na obrázku 81 obsahují informaci o tzv. Task a randomizační skupině. Volba Task je určena pro případy, kdy jsou vytvořeny dvě a více variant experimentu. Stimuly pro všechny varianty mohou být umístěny do stejného projektu a rozlišeny právě pomocí pole Task. Na začátku testování bude zobrazeno dialogové okno, ve kterém operátor zvolí, které části testu mají být prezentovány.

Pole Randomization Group slouží pro nastavení náhodné prezentace stimulů. Těchto randomizačních skupin lze v experimentu vytvořit několik, takže pokud má experiment například dvě části, budou stimuly sice zobrazeny náhodně, ale nebudou se míchat mezi skupinami. Velmi užitečná je rovněž volba „with previous“. Díky ní je zajištěno, že například fixační kříž bude zobrazen vždy před stimulem.

Po dokončení přípravy experimentu je nutné jej zamknout (přepínač Locked/Unlocked) v pravém horním rohu okna a uložit. Poté už může být spuštěno nahrávání dat. Před ostrým testováním je však vhodné vyzkoušet si experiment v pilotním testu tak, aby byly odhaleny případné chyby, nedostatky či bylo zjištěno, jak dlouho asi testování trvá.

### 5.5.2 Tvorba experimentu v Gazepoint Analysis

Tvorba experimentu v Gazepoint Analysis je specifická a liší se jak od tvorby v prostředí SMI Experiment Center, tak od tvorby v aplikaci OGAMA. Rozdíl je především v tom, že si program sám nevytváří složku projektu, ale je potřeba manuálně pro každý projekt (experiment) vytvořit složku. Po kliknutí na tlačítko „new project“ se zobrazí dialogové okno, kde je možné zvolit typ stimulu. Na výběr je text, obrázek, video, screen recording, web a web aggregate. Netradiční je, že není možné kombinovat jednotlivé typy stimulů mezi sebou v rámci jednoho projektu. Dohromady může být pouze text, obrázky a video. Pokud má být nahráván záznam obrazovky či web, musí být vytvořen nový experiment.



Obr. 82 Přidání nového stimulu v prostředí Gazepoint Analysis

Další nelogickou vlastností je to, že je sice možné nahrát několik stimulů (např. obrázků) najednou, ale ty budou poté mít v experimentu všechny

stejně označení (defaultně NewMedia) a při následné analýze je nebude možné odlišit (obr. 82). Z toho důvodu je nutné buď importovat stimuly jeden po druhém, nebo změnit jednotlivé názvy (pomocí ikony s nářadím). U každého stimulu může uživatel definovat dobu, po kterou bude zobrazen, a také zda má být zobrazen v náhodném pořadí. Náhodná prezentace skupin stimulů (otázka – fixační kříž – stimulus) není možná.

Pro použití textového stimulu je nutné vybrat soubor ve formátu TXT a ten bude automaticky zobrazen jako bílý text na černém pozadí. Stylování tohoto textu není možné. Pro zobrazení formátovaného textu je nutné jej prezentovat formou obrázku.

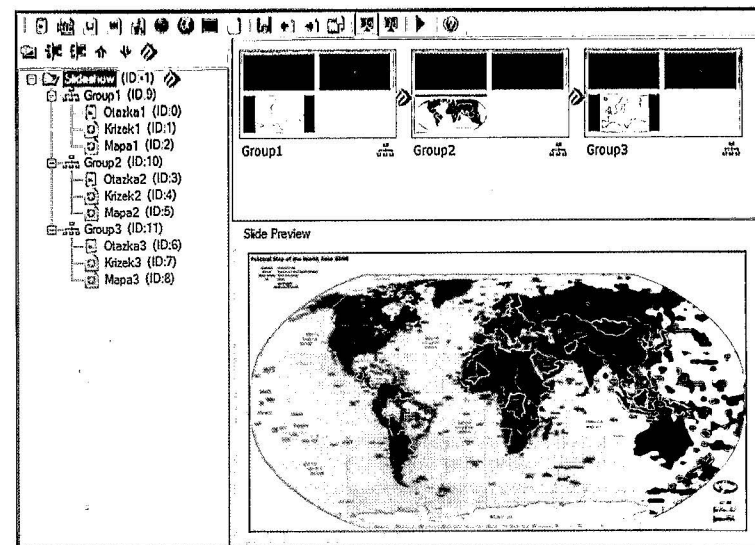
U stimulu typu Screen Recording se nenastavuje délka zobrazení a je nutné nahrávání manuálně ukončit (pomocí stisknutí „Stop Recording“ nebo pomocí klávesové zkratky Ctrl + Alt + S). Obdobně funguje i stimulus typu web. Jediný rozdíl oproti Screen Recording je to, že se automaticky zobrazí zvolená webová stránka.

Zajímavý je stimulus typu web aggregate, kdy jsou data více uživatelů přiřazena ke konkrétním webovým stránkám. V tomto případě je tedy možné nad každou navštívenou webovou stránkou zobrazit najednou data více uživatelů.

### 5.5.3 Tvorba experimentu v OGAMA

Aplikace OGAMA nabízí větší množství funkcí a nastavení než SMI Experiment Center, ale práce s ní není příliš intuitivní a často dochází k chybám. Z toho důvodu je užitečné vždy po několika krocích vyzkoušet, zda experiment funguje, a často jej ukládat. OGAMA umožňuje vytvářet stimuly typu text, obrázek, web a záznam obrazovky. Dále je možné vytvořit tzv. mixed slide kombinující několik typů dohromady. Zajímavostí je rovněž možnost vložení flashového videa. Pro kartografické účely je nejvhodnější použití obrázků, případně záznamu obrazovky. OGAMA však bohužel neumožňuje analýzu dat pomocí dynamických AOI.

Na obrázku 83 je ukázka jednoduchého vzorového experimentu, který obsahuje tři otázky vztahující se k mapám. Pomocí stimulu typu text byly vytvořeny texty otázek („Klikněte na...“). V OGAMA bohužel není možné kopírování slidů, takže například slide s fixačním křížkem je nutné vytvořit pro každou otázku zvlášť. Proto zde byl využit obrázek obsahující fixační kříž. Vytvoření stimulu s obrázkem je totiž rychlejší než použití textu, u kterého je nutné pokaždé specifikovat pozici. OGAMA umožňuje nastavit tzv. pre-slide fixation trial, což by bylo ideální právě pro fixační kříž. Ten by stačilo vytvořit jednou a dále už pouze u každého stimulu nastavit, ať je před ním zobrazen slide s fixačním křížem. Tato funkce ale bohužel nefunguje, pokud je nastaveno náhodné pořadí stimulů.



Obr. 83 Vzorový experiment vytvořený v aplikaci OGAMA

U vlastních obrázků map je nastaveno, že prezentace stimulu bude ukončena kliknutím levého tlačítka myši. Zde je velká výhoda OGAMA oproti např. SMI Experiment Center, neboť lze pro posun na další slide nastavit prakticky libovolnou akci – kliknutí myši, stisk klávesy, zobrazení po určitý časový interval. Tyto akce je navíc možné kombinovat, takže lze například nastavit, že prezentace stimulu bude ukončena po stisknutí pravé či levé šipky na klávesnici. Tyto akce jsou logovány, takže kromě vlastního posunu na další slide bude zaznamenána informace, kterou klávesu uživatel stiskl, čehož lze využít například pro vyjádření preference mezi dvěma obrázky. Informace o tom, kam uživatel v rámci stimulu klikl, je možné zjistit pomocí předem definované „oblasti zájmu“, označené zde jako „Target“.

Kliknutí do určité části stimulu či stisk určité klávesy lze rovněž využít pro přesměrování na další slide. Uživatel by si například na prvním stimulu zvolil, který typ map se mu více líbí, klikl by na něj a následovaly by už pouze stimuly v tomto stylu. U jednotlivých stimulů lze rovněž detailně nastavit pozadí. Kromě barvy to může být také obrázek či zvuk.

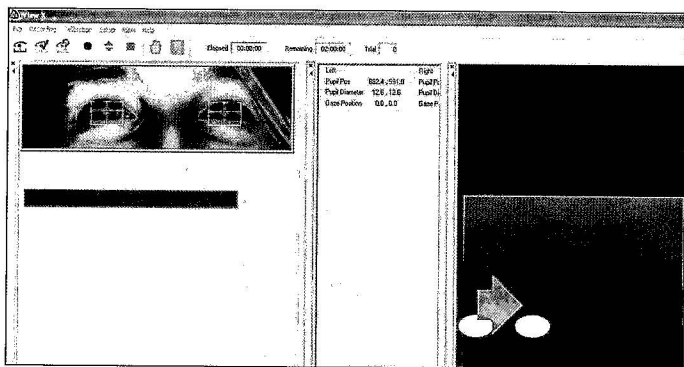
Pro nastavení náhodného pořadí stimulů bylo nutné nejprve vytvořit složky pro každou otázku (na obr. 83) označené jako Group 1–3. Pomocí navigačních tlačítek (vedle ikony pro novou složku) do nich byly následně umístěny stimuly. V dalším kroku byly vždy stimuly obsažené v jedné složce označeny (pomocí shift) a po stisku pravého tlačítka myši byla zvolena možnost „combine selected slides into one trial“. V posledním kroku stačí na složce Slideshow stisknout pravé tlačítko a zvolit „Shuffle subitems of selected item“. Zajímavá je možnost specifikovat, kolik stimulů bude

zobrazeno. Vzorový test obsahuje tři otázky, bylo by však možné nastavit že budou zobrazeny například pouze dvě. To by mohlo být užitečné například pro zkrácení celkové doby experimentu.

## 5.6 Záznam dat

Po dokončení tvorby experimentu může začít vlastní eye-tracking testování. Přesto je vhodné, aby každému „ostrému“ testování předcházelo pilotní testování na několika málo respondentech. Během něj lze odhadnout čas testování, zjistit nedostatky použitého experiment-designu a také zkontrolovat, zda se všechna data zaznamenávají. Rovněž je vhodné otestovat metody analýzy a vizualizace, které budou pro finální hodnocení naměřených dat použity. Obvyklá délka testování na katedře geoinformatiky nepřesahuje 15–20 minut. Ze zkušenosti bylo zjištěno, že při delších experimentech respondent neudrží pozornost. Přesto jsou pracoviště, kde eye-tracking testování trvá i více než dvě hodiny (např. Department of Teacher Education, University of Turku, Finsko). Pokud je taková délka experimentu nezbytná, je vhodné po určitém čase (např. 20 minut) zopakovat kalibraci.

Kalibrace je nezbytnou součástí každého experimentu. Ještě před spuštěním kalibrace je ale vhodné umístit respondenta před monitor tak, aby byl zaznamenávaný signál co nejlepší. Většina aplikací je k tomuto uzpůsobená, takže například SMI iViewX ukazuje kvalitu signálu pomocí barevného pruhu. Pokud je tento pruh zelený, sedí respondent správně a eye-tracker detekuje jeho oči. SMI iView X rovněž pomocí grafiky ukazuje, kterým směrem by se měl respondent posunout (přiblížit, oddálit), aby byl signál ještě lepší.



Obr. 84 Prostředí SMI iViewX zobrazující kvalitu signálu eye-trackera

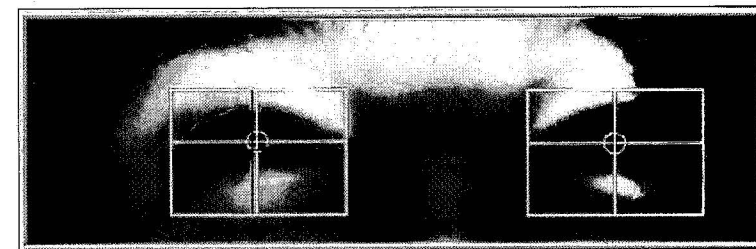
Na obrázku 84 je zobrazeno prostředí aplikace SMI iViewX. Dle barevného indikátoru (pruhu) je signál dostatečný (zelená barva), nicméně grafika na pravé straně ukazuje, že by se respondent měl posunout blíže k monitoru

a doprava. Hlava respondenta je totiž umístěna u okraje tzv. bounding boxu, a pokud by se v průběhu experimentu posunul ještě více doleva, mohlo by dojít ke ztrátě signálu.

Každý software obsahuje různé volby a možnosti kalibrace, přesto je tento proces ve všech případech velmi podobný. Úkolem respondenta je sledovat pohybující se tečku na monitoru. Počet míst, na kterých se tečka zastaví, neboli počet kalibračních bodů lze většinou v nastavení experimentu specifikovat. Ve většině aplikací je výstupem odchylka bodů, na které se respondent díval, od bodů, na které se měl dívat. Například v případě aplikace SMI Experiment Center lze zobrazit výsledek validace (validation), který ukazuje odchylku ve stupních. Je na výzkumníkovi, jak vysokou odchylku bude akceptovat, ale běžně se tolerují odchylky do 0,5° či 1° (Holmqvist a kol., 2011).

K problémům při kalibraci i záznamu dat může docházet kvůli různým vadám očí, jako je například astigmatismus, ale také kvůli brýlím s antireflexní vrstvou či díky řasence. Eye-tracker totiž detekuje zornici jako tmavý kruh a v některých případech se stává, že místo zornice detekuje řasenku. Dalším problémem mohou být povislá víčka u starších respondentů či obecně úzké oči. Na obrázku 85 je ukázka chybějícího korneálního odrazu u levého oka respondenta, tedy v pravé části obrázku.

Obecně se udává, že 5–10 % respondentů není možné z různých důvodů korektně zkalibrovat. S tímto faktem je třeba počítat před vlastní realizací eye-tracking experimentu a otestovat vyšší počet respondentů.



Obr. 85 Ukázka chybějící detekce korneálního odrazu u levého oka (v pravé části obrázku)

Po dokončení pilotního experimentu, ověření designu a srozumitelnosti pokládaných otázek může výzkumník přistoupit k vlastnímu průběhu experimentu a záznamu eye-tracking dat. Je vhodné, aby testování prováděl stále stejný výzkumník a aby všichni respondenti obdrželi stejné instrukce. Po dokončení testování je nezbytné nastavit parametry pro detekci fixací, ověřit kvalitu naměřených dat a případně data konvertovat do různých analytických nástrojů. Všechny uvedené procedury jsou popsány v následující kapitole.

## 7 ZPRACOVÁNÍ DAT

V této kapitole je popsáno několik metod užívaných pro zpracování eye-tracking dat. Vybrány byly metody vhodné pro použití v kartografických experimentech. Obecně existuje pro analýzu eye-tracking dat mnohem více metod, ty však uplatnění v kartografických experimentech nenajdou. Může se jednat například o analýzu regresí v experimentech zaměřených na čtení či analýzu průběhu sakád v klinických experimentech.

Vybrané metody slouží k vizualizaci či analýze eye-tracking dat. Hranice mezi těmito dvěma pojmy není ostrá, a proto se tyto pojmy mohou se prolínat. Jak uvádějí Andrienko a kol. (2012), při analýze pohybu očí jsou velmi často využívány metody tzv. vizuální analýzy (visual analytics) odvozené ze statistiky, strojového učení a dalších analytických disciplín. Tyto metody jsou speciálně navrženy pro použití lidskými analytiky. Hlavním cílem metod vizuální analýzy je pochopení pohybů očí a s tím související vhléd do základních kognitivních procesů.

Některé z níže popsaných metod mohou být zařazeny do kategorie prostého zobrazení naměřených dat, tedy vizualizace. Na základě vizuální analýzy tohoto výstupu však výzkumník může analyzovat chování respondenta, a dozvědět se tak něco o kognitivních procesech, které při čtení stimulu u respondenta probíhaly.

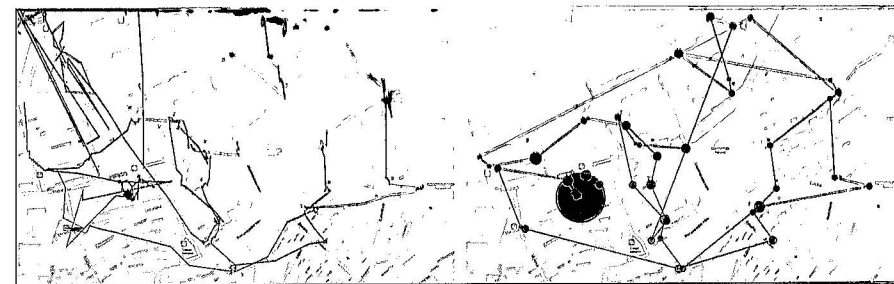
### 7.1 Vizualizace trajektorií

Mezi nejjednodušší metody vizuální analýzy patří prosté vykreslení trajektorie oka nad studovaným stimulem. K tomu se nejčastěji používají dvě metody. První z nich je GazeReplay, kdy je trajektorie oka zobrazena formou animace a je zobrazena pouze několikavteřinová část trajektorie. Druhou metodou je scanpath, někdy též označovaný jako GazePlot. V tomto případě je výstupem statický obrázek zobrazující trajektorii oka za celou dobu trvání trialu (pozorování jednoho stimulu uživatelem).

**GazeReplay** je metoda vizualizace eye-tracking dat, kdy jsou informace o fixacích a sakádách zobrazovány dynamicky nad sledovaným obrazem měnícím se v čase. GazeReplay lze využít při kvalitativním vyhodnocení uživatelské percepce studovaného stimulu. GazeReplay je nejpřesnější metodou pro analýzu eye-tracking dat, protože nedochází k žádné interpolaci ani generalizaci (Nielsen a Pernice, 2010). GazeReplay ukazuje přesně pozici, kam se uživatel díval. Analýza těchto výsledků je ale velice složitá. Pomocí GazeReplay je možné přehrát videozáznam práce uživatele s dokumentem, přičemž součástí obrazu je bod, na který se uživatel v průběhu videozáznamu soustředil. Vhodné je tento přístup zkombinovat s metodou think-aloud, což umožní analyzovat pohyby uživatelových očí v souvislosti s konkrétní činností.

Velmi cennou metodou pro zobrazování dat z eye-trackingu je tzv. **scanpath**. Podle Holmqvista (2011) byl tento termín poprvé použit v 70. letech 20. století. Scanpath bývá také označován jako GazePlot, Fixation Track nebo Eye-Movement Pattern. Jedná se o trajektorie sakád spojující pozice fixací, zobrazené přes studovaný obraz, který slouží jako podklad. Scanpath zobrazuje fixace jako kruhy (případně kříže) o různé velikosti (jejich poloměr odpovídá délce fixací) a sakády jako linie, které tyto kruhy spojují (Raiha a kol., 2005). Omezení této metody nastává v případě zobrazování většího množství dat, kdy díky překryvům jednotlivých fixací není možné vizuálně poznat jejich počet. Scanpath lze, podobně jako výše zmíněný GazeReplay, využít zejména při kvalitativním hodnocení eye-tracking experimentu. Vizuálním hodnocením scanpath se zabýval například Buswell (1935), který hodnotil uživatelské vnímání obrazů na základě pořadí a pozice jednotlivých fixací a vyhodnocením míst, na která se uživatel díval (fixoval) a na která ne. Scanpath je možné využít také při ověření kvality naměřených dat, získání představy o rozmístění fixací nebo jako ilustrační obrázek popisující chování uživatele v rámci stimulu.

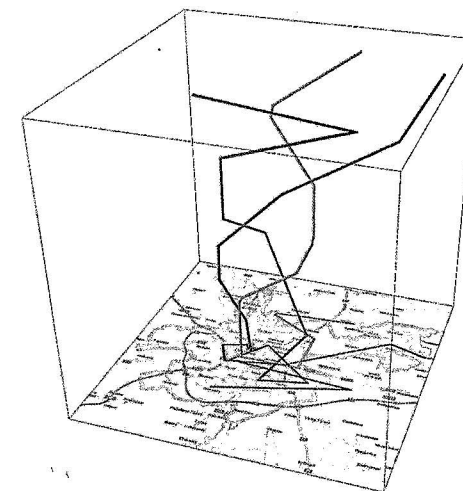
V obou případech je možné zobrazit jak původní naměřená data, tak i z nich odvozené fixace a sakády. Postup tohoto odvození je popsán v předchozí kapitole. Zobrazení fixací a sakád na místo původních naměřených (raw) dat je výhodné ze dvou důvodů. Prvním z nich je větší přehlednost celé vizualizace, druhým pak filtrace chyb měření. K těm může docházet v případě, že zařízení ztratí na několik málo milisekund signál. V tomto případě se pak místo skutečné souřadnice pohledu oka zapíše hodnota (0;0). Při zobrazení surových dat se pak části trajektorie oka zobrazují v levém horním rohu stimulu, jak je to patrné z obrázku 106.



Obr. 106 Zobrazení naměřených (raw) dat (vlevo) a vizualizace fixací a sakád formou scanpath (vpravo)

Při zobrazení scanpath v prostředí SMI BeGaze může uživatel specifikovat barvu (pomocí záložky Dashboard) trajektorií jednotlivých respondentů. Dále je možné nastavit, zda budou fixace zobrazeny jako kruhy či kříže (crosshair), zda budou všechny fixace stejně velké či bude velikost kruhu (kříže) odpovídat délce fixace. Zobrazit lze také čísla označující pořadí fixací. Zajímavou možností je barevné rozdělení scanpath dle časových intervalů. Bohužel je ale možné zadat pouze dva lomové body.

Možnosti vizualizace scanpath v OGAMA jsou velmi podobné. Na rozdíl od SMI BeGaze ale OGAMA v rámci scanpath modulu umožňuje vygenerovat sekvenci navštívených oblastí zájmu. Tato data lze poté použít pro exaktní výpočet podobnosti trajektorií pohybu očí několika respondentů pomocí nástroje ScanGraph, který je detailně popsán v kapitole 7.4.



Obr. 107 Koncept Space-Time-Cube

Eye-tracking data lze zobrazit v prostoru pomocí tzv. Space-Time-Cube (Hägerstrand, 1970). Scanpath je v tomto případě zobrazen uvnitř krychle, přičemž základna krychle (osa X a Y) slouží k zobrazení stimulu a ve směru osy Z je zobrazen čas. Hlavní předností této vizualizace je zobrazení prostorové i časové složky najednou. Jak je patrné z obrázku 107, pokud v čase nedochází ke změně polohy pohledu, je linie kolmá k základně krychle. Čím pozvolnější je linie mezi dvěma lomovými body, tím rychleji došlo ke změně polohy pohledu.

Vizualizací eye-tracking dat pomocí Space-Time-Cube se ve své disertační práci zabývala Kveladze (2015), která využívala nástroj ILWIS. Pomocí Space-Time-Cube lze eye-tracking data zobrazit i v aplikaci V-Analytics. Tento postup je popsán v kapitole 7.5.

## 7.2 Attention mapy

Attention mapy jsou nástrojem k vizualizaci kvantitativních charakteristik pohledu uživatele. Attention mapy bývají často označovány jako heat mapy, což však není nejvhodnější, a to zejména v kontextu kartografie, kde je heat mapa považována za teplotní mapu, tedy mapu teploty. Z attention map je patrné, které oblasti sledovaného obrazu uživatel zkoumá pohledem více a kterým naopak nevěnuje pozornost. Attention mapy jsou v eye-trackingu velice výhodné pro vytvoření rychlého přehledu, na které části dokumentu se uživatelé soustředí a které je vhodné hlouběji analyzovat.

Attention mapy se vytvářejí dvěma zdánlivě rozdílnými, ale přesto velice podobnými principy (Holmqvist a kol., 2011). Jedná se o tzv. Gridded AOI a topologický (gaussovský) povrch. Oba druhy attention map je možné vytvářet z prvotních (raw) i klasifikovaných dat (fixací).

První zmíněná metoda rozdělí prostor (stimulus) na pravoúhloú síť oblastí zájmu (AOI). Zapsáním hodnoty určité eye-tracking metriky (Dwell Time, Fixation Count) a obarvením výsledku dle této hodnoty vzniká attention mapa. Kromě výše zmíněných a velice často využívaných metrik může být vizualizována například metrika Entry time – čas, kdy byla v oblasti poprvé zaznamenána fixace. Takto je možné získat přehled o tom, které části stimulu uživatelé navštívili na začátku trialu a které později. Navzdory této univerzálnosti nejsou attention mapy vytvořené tímto principem těmi nejčastěji využívanými. Podrobněji jsou Gridded AOI popsány v kapitole 7.3.

V případě Gridded AOI attention map jsou hranice mezi buňkami ostré. Oproti tomu u Gaussových attention map jsou přechody plynulé. Gaussovy attention mapy jsou vytvářeny postupným přidáváním hodnoty délky fixací k jednotlivým pixelům stimulu. Protože je takto ohodnoceno pouze relativně malé množství pixelů, jsou tato data proložena Gaussovou funkcí.

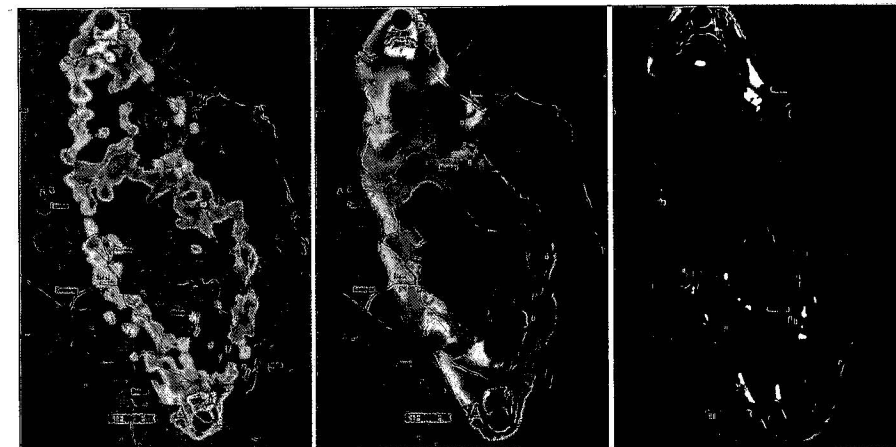
Tato funkce je definována jako:

$$G(x, y) = \exp\left(-\frac{(x - y_i)^2 + (y - y_i)^2}{2\sigma^2}\right)$$

kde  $x_i$  a  $y_i$  jsou souřadnice středu fixace a hodnoty  $x$  a  $y$  odpovídají rozlišení stimulu.

Výsledná attention mapa vzniká aplikací Gaussovy funkce pro každý střed fixace a následným nahrazením všech těchto dílčích funkcí jedinou Gaussovou funkcí (upraveno podle Holmqvista a kol., 2011). Barva attention map ve většině komerčních aplikací reprezentuje čas v milisekundách.

Při vytváření Attention map je velice důležité dbát na nastavení hodnoty  $\sigma$  (sigma), označované též jako kernel width, která odpovídá velikosti shlazení výsledné funkce. Při nastavení nízkých hodnot bude výsledná mapa pokrývat pouze oblast v bezprostředním okolí středů zaznamenaných fixací. Oproti tomu vysoká hodnota sigma způsobí větší rozšíření attention mapy. Výsledná vizualizace bude výraznější, ale také méně přesná. Vysoké hodnoty sledovaného jevu budou i v oblasti, kde žádné fixace zaznamenány nebyly. Vliv rozdílného nastavení hodnoty kernel width v programu OGAMA je znázorněn na obrázku 108.

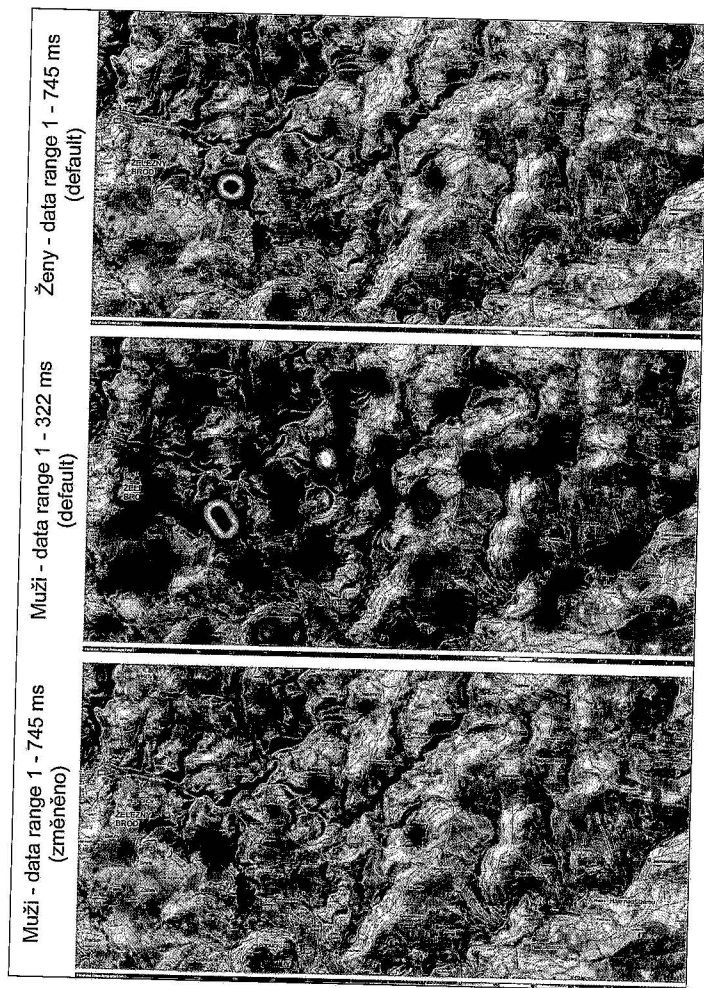


Obr. 108 Vliv nastavení hodnoty kernel width u attention mapy v programu OGAMA. Vlevo je nastavena hodnota 50, uprostřed 100 a vpravo 200.

Attention mapy bývají často používány pro srovnání rozdílného chování respondentů nad několika stimuly nebo rozdílného chování skupin respondentů. Závažným problémem je v tomto případě různé nastavení rozsahu dat (data range), pro který byla mapa generována. Na základě tohoto rozsahu je vytvářena barevná stupnice. V případě rozdílného nastavení nejsou výstupy



vzájemně porovnatelné. Příkladem může být obrázek 109. První dva obrázky ukazují attention mapu vytvořenou pro skupinu žen (nahore) a mužů (uprostřed), kdy byly ponechány defaultní hodnoty rozsahu dat z programu SMI BeGaze. Porovnání těchto dvou obrázků by vedlo k závěru, že muži si prohlédli mnohem více míst v mapě. Důležité však je, že defaultní hodnota rozsahu dat je pro obě skupiny respondentů velmi odlišná (1–745 ms pro ženy a 1–322 ms pro muže). Při změně rozsahu dat pro skupinu mužů na stejnou hodnotu, jako měly ženy, se výsledná attention mapa zásadně změní a rozdíl mezi oběma skupinami už není zásadní (obr. 109 dole).

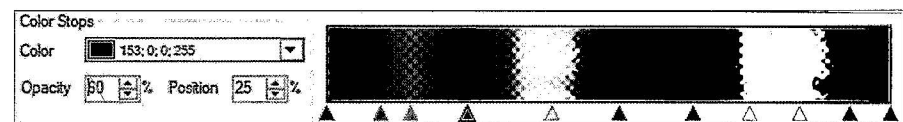


Obr. 109 Vliv nastavení rozsahu dat (data range) na výslednou attention mapu. První dva obrázky ukazují attention mapu vytvořenou s použitím defaultních hodnot rozsahu dat z programu SMI BeGaze. Obrázek dole ukazuje attention mapu pro skupinu mužů s opravenou hodnotou rozsahu dat.

V programu SMI BeGaze lze v nastavení nástroje Heat Map kromě hodnoty sigma (Kernel width) a rozsahu dat (Data range) upravit také průhlednost (Opacity). Nevýhodou jsou však omezené možnosti výběru barevné škály. Vybrat lze ze dvou přednastavených variant (modrá – zelená – červená a zelená – žlutá – červená). Dále si může uživatel vytvořit škálu vlastní, nicméně opět pouze pomocí tří barev (začátek – střed – konec). Výsledné attention mapy jsou však vždy velmi nevýrazné.

SMI BeGaze kromě nástroje HeatMap obsahuje také nástroj Focus Map, který funguje tak, že celý stimulus je překryt jednobarevnou vrstvou a transparentní jsou pouze ta místa, na kterých byly zaznamenány fixace. V nastavení lze stejně jako v případě heat map specifikovat kernel width, rozsah dat a průhlednost.

V aplikaci OGAMA lze attention mapy vytvářet pomocí Attention map modulu. Výhodou oproti SMI BeGaze je možnost nastavení vlastní barevné stupnice obsahující libovolný počet barev (obr. 110) a také možnost přiřazení vyšší váhy delším fixacím. Zajímavá je také možnost vykreslit attention mapu pouze pro konkrétní (první, druhou, třetí...) fixaci či generování attention mapy na základě pohybu či kliků myši. Zásadním problémem v prostředí OGAMA je však nemožnost definovat vlastní rozsah dat. Attention mapy vytvořené v OGAMA tedy není možné použít pro porovnání různých stimulů či skupin respondentů.



Obr. 110 Nastavení barevné škály attention map v prostředí aplikace OGAMA

Holmqvist a kol. (2011) uvádějí několik základních doporučení, na která je důležité při používání attention map pamatovat.

- Attention mapy ukazují prostorové rozložení fixací (či raw dat) a nic jiného, nelze z nich tedy usuzovat, proč se na dané místo stimulu respondenti dívali.
- Attention mapy zobrazují agregovaná data za celou délku sledování stimulu a za více participantů.
- Attention mapy jsou vhodným ilustračním nástrojem, jejich vypovídací hodnota však není vysoká a měly by být publikovány až po důkladném zvážení.
- Neexistuje žádné přesné doporučení ohledně nastavení  $\sigma$ . Uvádí se, že hodnota okolo  $2^\circ$  vizuálního úhlu (cca 85 px) přibližně odpovídá oblasti ostrého vidění (světlo dopadající na foveu).

- Při tvorbě attention map pro různé skupiny respondentů či stimulů je nezbytné použít vždy stejné nastavení, jinak nebude možné výsledky porovnat.
- Při publikování attention map je vždy nezbytné popsat typ eye-trackeru, na kterém byla data měřena, jeho frekvenci, nastavení  $\sigma$ , rozsah dat, pro který byla attention mapa vytvářena, a podrobnosti o algoritmu pro detekci fixací.
- Při použití fixačního kříže je důležité při interpretaci výsledků attention mapy pamatovat na fakt, že první fixace byla zaznamenána uprostřed obrazovky, případně první fixaci odstranit.

### 7.3 AOI

Další možností analýzy eye-tracking dat je využití oblastí zájmu (Areas of Interest – AOI). Oblasti zájmu jsou regiony vyznačené na stimulu, pro které je zjišťováno, jak respondenta zaujaly, kolik fixací bylo v konkrétních oblastech zájmu zaznamenáno, jaké bylo pořadí navštívených oblastí atd. Oblasti zájmu není nutné definovat před vlastním testováním, protože tvorba AOI probíhá až po vlastním naměření dat. Tvorba AOI v programech SMI BeGaze a OGAMA je velmi podobná, přesto se najdou určité rozdíly. Na základě nich se může uživatel rozhodnout, ve které aplikaci bude data zpracovávat. Obě aplikace umožňují základní kreslení pomocí nástroje obdélník, elipsa a vlastní tvar. Aplikace OGAMA dále umožňuje vytvoření pravidelné mřížky (až po 26×26 buněk).

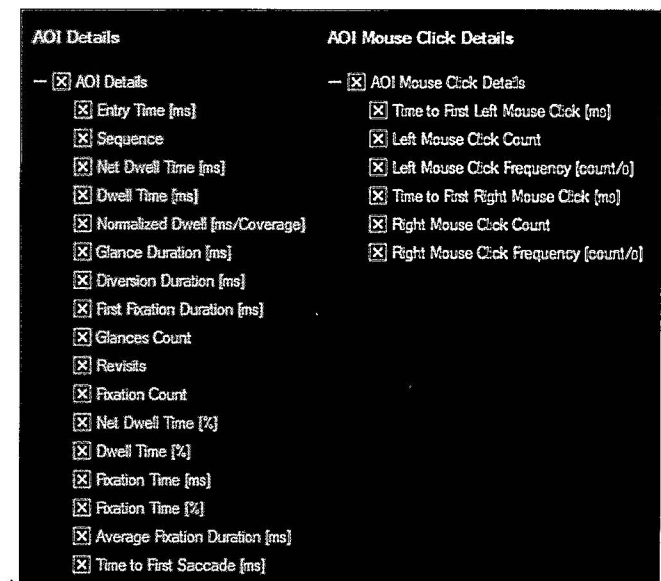
Pokud uživatel zakreslí překrývající se oblasti zájmu, musí počítat s tím, že fixace zaznamenaná v překrývající se oblasti bude zaznamenaná k oběma AOI. V aplikaci SMI BeGaze je možné nastavit pořadí oblastí zájmu. Toto je však použito pouze pro metody vizualizace Sequence Chart a Binning Chart (viz dále).

Na rozdíl od programu OGAMA umožňuje SMI BeGaze práci s dynamickými AOI. Uživatel tak může zakreslovat oblasti zájmu do záznamu obrazovky, videozáznamu či záznamu z eye-tracking brýlí. Tvorba dynamických oblastí zájmu probíhá obdobně jako v případě statických AOI. Rozdíl je ten, že se uživatel pohybuje na časové ose a při každé změně polohy sledovaného objektu (kolem kterého chce zakreslit AOI) přesune AOI na správné místo. Aplikace SMI BeGaze při každé změně automaticky vytvoří tzv. key frame. Pokud se mezi dvěma key frames změni poloha AOI, aplikace ji plynule přesune. V některých případech je toto chování nežádoucí, proto je vhodné manuálního umístění key frame. Tato volba je dostupná po stisku pravého tlačítka myši na označené oblasti zájmu. Zde je dále možné oblast zájmu zneviditelnit. Při vyšším počtu oblastí zájmu je tento postup velmi časově náročný, a to zejména v případě, že má každý respondent unikátní videozáznam (data pořízená jako screen-recording nebo data z eye-tracking brýlí).

Významný rozdíl mezi oběma aplikacemi je ve způsobu kopírování AOI mezi více stimulů a především v exportu eye-tracking metrik, které mohou být vypočítány pro každou oblast zájmu zvlášť. Podrobnosti o eye-tracking metrikách jsou popsány v kapitole 7.8.1.

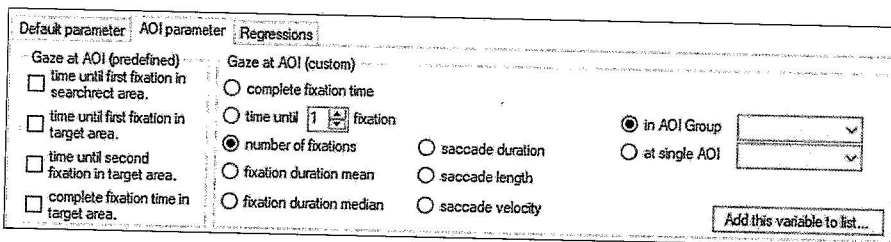
Aplikace OGAMA umožňuje kopírování AOI mezi stimuly pomocí klávesových zkratk Ctrl+C a Ctrl+V. Oproti tomu v SMI BeGaze je nutné nejprve sadu AOI uložit do XML souboru a poté u dalších stimulů tento soubor načíst. Druhou možností je nastavit ve vlastnostech jednotlivých oblastí zájmu rozsah (scope) na hodnotu global. Tato oblast zájmu bude poté zobrazena u všech stimulů v experimentu. Oblasti zájmu je možné v obou programech seskupovat do skupin. Toho lze využít, pokud jsou cílem studia jak konkrétní oblasti zájmu, tak i jejich skupiny. Příkladem může být legenda mapy rozdělená na více částí. Při analýze je možné studovat jak eye-tracking metriky pro jednotlivé části, tak legendu jako celek. V SMI BeGaze k tomu slouží volba group, do které může uživatel přímo napsat název skupiny. V programu OGAMA je seskupování o něco složitější, protože je nejprve nutné tyto skupiny vytvořit prostřednictvím Statistics modulu – záložka Shape Groups.

Export statistických údajů (eye-tracking metrik) pro definované oblasti zájmu probíhá v SMI BeGaze pomocí funkce Metrics export a záložky AOI Statistics, a to pro jednotlivé AOI nebo pro jejich skupiny. Přehled eye-tracking metrik, které je možné v programu SMI BeGaze pro oblasti zájmu exportovat, je uveden v obrázku 111.



Obr. 111 Eye-tracking metriky pro oblasti zájmu v programu SMI BeGaze

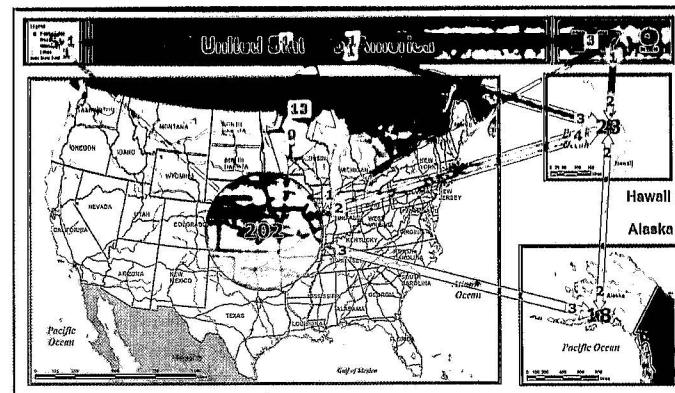
V programu OGAMA se eye-tracking metriky pro oblasti zájmu exportují pomocí Statistics modulu – Gaze parameters – AOI parameter. Jak je vidět z obrázku 112, seznam dostupných eye-tracking metrik je kratší než v případě SMI BeGaze, nicméně většina často používaných metrik je zde obsažena. Při exportu statistických dat si uživatel zvolí vybranou eye-tracking metriku a dále vybere, pro kterou AOI či skupinu AOI má být tato metrika vypočítána. Následně svoji volbu potvrdí stiskem tlačítka Add this variable to the list. V případě vyššího počtu oblastí zájmu v experimentu může být tento postup časově náročný.



Obr. 112 Eye-tracking metriky pro oblasti zájmu v programu SMI BeGaze

Obě popisované aplikace umožňují export tzv. Transition Matrix, tedy matice reprezentující přechod pohledu respondenta mezi jednotlivými oblastmi zájmu. V této matici jsou ve sloupcích i řádcích vypsány oblasti zájmu a hodnota v buňkách matice udává, kolikrát se pohled respondenta přesunul z jedné AOI do druhé. Tato data lze statisticky vyhodnotit nebo vizualizovat. V SMI BeGaze je matice dostupná v nabídce Metrics export – Specialized statistics – Transition Matrix. V programu OGAMA pak ve Statistics modulu – AOI Transitions. Rozdíl mezi oběma aplikacemi ve výsledné matici je minimální. OGAMA v matici uvádí i počet přesunů pohledu v rámci jedné oblasti zájmu, zatímco SMI BeGaze mezi stejnými AOI uvádí nuly. Obě aplikace znázorňují v řádcích oblasti zájmu, ze kterých přesuny vycházejí, a ve sloupcích oblasti, kam míří.

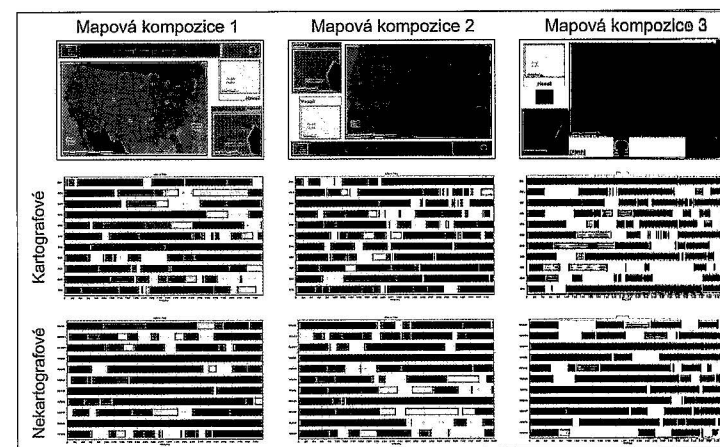
Program OGAMA umožňuje znázornit tyto přesuny pohledu pomocí šipek, a to jak v absolutních číslech, tak i relativně. Další zajímavou vizualizační funkcí je zobrazení kruhů reprezentujících počet fixací, celkový čas fixací či průměrnou délku fixace. Obě tyto metody jsou dostupné přímo v AOI modulu programu OGAMA a lze je zobrazit najednou tak, jak je znázorněno na obrázku 113.



Obr. 113 Vizualizace počtu přesunů pohledu mezi oblastmi zájmu a celkového počtu fixací v jednotlivých AOI v prostředí programu OGAMA

Příkladem dalších metod vizualizace respondentova chování využívající oblasti zájmu jsou nástroje Sequence Chart a Binning Chart obsažené v softwaru SMI BeGaze.

**Sequence Chart** zobrazuje časovou posloupnost navštívených oblastí zájmu. Eye-tracking data jednotlivých respondentů jsou znázorněna pomocí barevných pruhů, přičemž barva jednotlivých pruhů reprezentuje jednotlivé oblasti zájmu. Ze Sequence Chartu je tedy patrné, jaké bylo pořadí navštívených oblastí zájmu, jak dlouho v nich respondenti spočinuli pohledem a jestli se do některé AOI dívali opakovaně.



Obr. 114 Sequence Chart dvou skupin respondentů (kartografové a nekartografové) pro tři různé mapové kompozice. Skupina respondentů s kartografickým vzděláním se mnohem déle dívala na oblast zájmu představující titul mapy.