

TAMPEREEN YLIOPISTO

Sami Serola

**Opiskelijoiden tiedontarpeet ja viitteiden hyödyllisyys
tutkimussuunnitelman laatimista varten tehdyissä tiedonhauissa**

Informaatiotutkimuksen pro gradu –tutkielma

Tampere 2001

Tampereen yliopisto

Informaatiotutkimuksen laitos

SEROLA, SAMI: Opiskelijoiden tiedontarpeet ja viitteiden hyödyllisyys tutkimussuunnitelman laatimista varten tehdyissä tiedonhauissa

Pro gradu –tutkielma, 49 s., 10 liites.

Informaatiotutkimus

Marraskuu 2001-11-27

Käyttäjakeskeisissä tiedonhaku tutkimuksissa on yleensä tarkasteltu dokumenttien tai viitteiden hyödyllisyyttä riippumatta tehtävän vaiheesta. Tämän tutkimuksen tavoitteena oli selvittää, minkä tyyppistä tietoa käyttäjät odottivat löytävänsä ja minkä tyyppistä tietoa he arvioivat löytäneensä tehtävän eri vaiheissa. Lisäksi selvitettiin miten tehtävän suorituksen vaihe oli yhteydessä käyttäjien arvioihin viitteiden hyödyllisyydestä, aiheenmukaisuudesta ja tiedonhakujen onnistuneisuudesta yleensä.

Tutkimuksen kohteena oli 22 psykologian opiskelijaa, joiden tehtävänä oli laatia tutkimussuunnitelma suppeaa empiiristä tutkimusta varten vuoden 2000 syys- marraskuussa. Opiskelijoiden tekemiä viitetietokantahakuja havainnointiin tehtävän alkuvaiheessa ja sen loppupuolella. Aineistoa kerättiin muun muassa kyselylomakkeiden, hakutallenteiden, strukturoitujen teemahaastattelujen ja opiskelijoiden pitämien hakupäiväkirjojen avulla.

Tehtävän alussa psykologian opiskelijat odottivat löytävänsä lähinnä yleistä taustatietoa, mutta viitteiden perusteella taustatietoa ei löydetty odotetussa määrin. Viitteiden arvioitiin antavan omalle työlle lähinnä yleistä taustatietoa, empiriatietoa ja odotettua enemmän myös yksityiskohtaista tietoa. Tehtävän lopussa psykologian opiskelijat tarvitsivat ensisijaisesti menetelmä- ja empiriatietoa. Opiskelijat löysivät eniten yksityiskohtaista tietoa ja empiriatietoa, mutta yleisesti ottaen mitään kaivatun tyyppistä tietoa ei löydetty odotetussa määrin toisella hakukerralla.

Tehtävän alussa psykologian opiskelijat löysivät enemmän hyödyllisiä viitteitä kuin tehtävän lopussa. Molemmilla hakukerroilla suurin osa viitteistä arvioitiin mahdollisesti hyödyllisiksi. Opiskelijoiden kriittisyys hyödyllisyyden suhteen kuitenkin lisääntyi tehtävän aikana, sillä tehtävän loppupuolella hyväksyttiin vähemmän viitteitä täysin hyödyllisiksi. Molemmilla hakukerroilla opiskelijat pitivät hakuja kattavina ja he olivat tyytyväisiä saamiinsa viitteisiin, kun he löysivät paljon täysin hyödyllisiä ja aiheenmukaisia viitteitä. Ensimmäisellä hakukerralla osa opiskelijoista oli kattavuudesta riippumatta tyytyväisiä saamiinsa viitteisiin, kun he löysivät paljon viitteitä ja kun viitteiden arvioitiin antavan yksityiskohtaista tietoa. Toisella hakukerralla osa opiskelijoista oli kattavuudesta riippumatta tyytyväisiä erityisesti mahdollisimman aiheenmukaisiin viitteisiin.

Tutkimustulosten perusteella hakujärjestelmien käyttöliittymien tulisi tarjota helppokäyttöisempiä apuvälineitä tehtävän eri vaiheissa oleville käyttäjille. Tiedonhaun alussa olevat voisivat esimerkiksi hyötyä linkitetystä aihehakemistosta. Tehtävässä pidemmällä edenneille järjestelmä voisi puolestaan automaattisesti tarjota käyttäjän hakutermejä vastaavia asiasanoja.

Sisällys

Esipuhe

1. Johdanto	1
2. Teoreettinen tausta ja keskeiset käsitteet	2
2.1. <i>Aihetietämyksen muutos tiedon hankinnan ja tehtävän suorituksen eri vaiheissa</i>	2
2.2. <i>Relevanssin käsite informaatiotutkimuksessa</i>	4
2.3. <i>Käyttäjakeskeisen relevanssin eri asteet</i>	6
2.4. <i>Informaatio-objektien valintaan vaikuttavat kriteerit</i>	6
2.5. <i>Haetun tiedon anti hyödyllisyyden mittana</i>	7
2.6. <i>Tiedonhaun yleinen onnistuneisuus käyttäjän näkökulmasta</i>	9
3. Aikaisempi tutkimus	10
3.1. <i>Relevanssin asteet suhteessa aihetietämyksen määrään</i>	10
3.2. <i>Tiedonhankinnan prosessimalli pitkäaikaistutkimuksen perustana</i>	12
3.3. <i>Aihetietämyksen lisääntyminen tiedonhankinnan eri vaiheissa</i>	12
3.4. <i>Tiedon tarve vaativan tehtävän suorituksen eri vaiheissa</i>	13
4. Tutkimuskysymykset ja hypoteesit	15
4.1. <i>Tiedon odotettu tarve arvioituna tietotyypeittäin</i>	16
4.2. <i>Viitteiden hyödyllisyysarvioiden muutos</i>	18
4.3. <i>Viitteiden tietotyyppiarvioiden muutos</i>	18
4.4. <i>Haun onnistuminen käyttäjän näkökulmasta</i>	19
5. Tutkimuksen kulku ja käytetyt menetelmät	20
5.1. <i>Aineistonkeruumenetelmät</i>	20
5.2. <i>Haastattelujen ja tiedonhakujen kulku</i>	21
5.3. <i>Datan käsittely tarkastelua varten</i>	22
5.4. <i>Tilastollisten analyysimenetelmien valinta</i>	23
6. Tulokset	26
6.1. <i>Tietotyyppien odotettu tarve</i>	26
6.2. <i>Tietotyypeittäin arvioitu viitteiden anti</i>	29
6.3. <i>Tietotyyppiodotusten ja viitteiden tietotyyppiarvioiden vastaavuus</i>	31
6.4. <i>Löytyneiden viitteiden hyödyllisyys</i>	32
6.5. <i>Viitteiden aiheenmukaisuuden piirteet</i>	34
6.6. <i>Haun onnistuneisuus yleensä</i>	36
7. Tulosten tarkastelu	42
7.1. <i>Hyödyllisyysarviot eri käyttäjäryhmien välillä</i>	43
7.2. <i>Aiheenmukaisuuden ja hyödyllisyyden piirteet</i>	44
7.3. <i>Haun onnistuneisuus kaikkiaan</i>	45
7.4. <i>Tutkimuksen anti</i>	46
Lähteet	48
Liitteet	

Lista kuvioista ja taulukoista

Kuviot

1. Kuhlthaun tiedonhankinnan prosessimallin vaiheet, ja niiden aikana ilmenevät tuntemukset, ajatukset, toiminnot ja tehtävät	3
2. Tutkimuksessa tarkasteltujen muuttujien väliset suhteet	15
3. Tutkimuksessa käytetyt tietotyyppi- jaottelut	17
4. Teoriatiedon odotettu anti tehtävän alussa ja lopussa	24
5. Aiheen täsmennystä auttavan tiedon odotettu anti mitattuna tehtävän alussa ja lopussa	24
6. Tutkimuksessa käytetty kaava osittaiskorrelaation laskemiseksi	26
7. Aiheenvalintaa auttavien ideoiden odotettu anti ensimmäisellä haastattelukerralla.....	28
8. Saatujen viitteiden kokonaistyytyväisyyden selittyminen viitteen piirteen perusteella, kun kattavuus on vakioitu.....	39
9. Haun onnistuneisuusarvioiden keskinäinen riippuvuus sekä riippuvuus viitteiden tietotyyppiannista, aiheenmukaisuudesta, määrästä ja täysin hyödyllisten viitteiden osuudesta ensimmäisellä hakukerralla	40
10. Haun onnistuneisuusarvioiden keskinäinen riippuvuus sekä riippuvuus viitteiden tietotyyppiannista, aiheenmukaisuudesta, määrästä ja täysin hyödyllisten viitteiden osuudesta toisella hakukerralla	41

Taulukot

1. Tietotyyppien odotettu tarve ennen tiedonhakuja: keskiarvot hakukerroittain per henkilö sekä eri vaiheitten välisen eron tilastollinen merkitsevyys.....	27
2. Löytyneiden hyödyllisten viitteiden arvioidun tietotyyppiannin ja aiheenmukaisuuden keskiarvot laskettuna viitteille.....	29
3. Viitteiden tietotyyppiannin faktorit eri hakukerroilla	30
4. Tietotyyppiodotusten sekä viitteiden arvioidun annin keskiarvot per henkilö sekä niiden välisen eron tilastollinen merkitsevyys vaiheittain	31
5. Viitteiden määrien keskiarvot ja prosentuaalinen osuus hyödyllisyysasteittain eri hakukerroilla sekä erojen tilastollinen merkitsevyys vaiheitten välillä.....	32
6. Korrelaatiot viitteiden arvioidun hyödyllisyyden, tietotyyppiannin ja aiheenmukaisuuden välillä.....	33
7. Osuvasti aiheenmukaisten viitteiden tietotyyppiarvioiden ja hyödyllisyysarvion keskiarvot, sekä erojen tilastollinen merkitsevyys.....	35
8. Löytyneiden hyödyllisten viitteiden aiheenmukaisuuden ja tietotyyppiannin korrelaatiot eri hakukerroilla	35
9. Haun kattavuusarvio ja tyytyväisyys saatuihin viitteisiin kokonaisuudessaan. Keskiarvot per henkilö hakukerroittain sekä erojen tilastollinen merkitsevyys vaiheitten välillä.....	36
10. Löytyneitten hyödyllisten viitteiden piirteiden korrelaatiot suhteessa viitteiden määrään ja haun onnistuneisuusarvioihin laskettuna henkilöittäin	37
11. Osittaiskorrelaatiot viitteiden piirteiden ja saatujen viitteiden kokonaistyytyväisyyden välillä, kun arvio haun kattavuudesta on vakioitu.....	39
12. Haun onnistuneisuusarvioiden korrelaatiot suhteessa tietotyyppiodotusten ja viitteiden tietotyyppiannin väliseen eroon	41

Esipuhe

Tämä työ on tehty professori Pertti Vakkarin oivallisessa ohjauksessa ja Mikko Pennasen hyvien kommenttien avustuksella. Työn ovat mahdollistaneet myös Tampereen yliopiston informaatiotutkimuksen laitoksen rahoitus sekä tutkimukseen osallistuneiden 22:n psykologian opiskelijan aktiivisuus.

Tutkimus on auttanut minua perehtymään perusteellisesti tieteelliseen tutkimukseen. Olen muun muassa varmistunut siitä, että myös pienten populaatioiden ja järjestysasteikollisten muuttujien tarkastelu on mahdollista oikein valittujen tilastollisten menetelmien avulla. Työni on tosin liiaksikin painottunut kvantitatiivisten menetelmien soveltamiseen, jolloin kvalitatiivinen tarkastelu on jäänyt kovin vähälle huomiolle. Päädyin kuitenkin tähän ratkaisuun siksi, että katsoin tarpeelliseksi perehtyä ensin yksityiskohtaisemmin kvantitatiivisiin menetelmiin. Työn edistyessä kävi myös yhä ilmeisemmäksi, etten pysyisi aikataulussa, jos yrittäisin yhtä yksityiskohtaisesti paneutua aineistoon myös kvalitatiivisesti. Monta mielenkiintoista havaintoa on jäänyt tämän työn ulkopuolelle, joista toivon voivani raportoida myöhemmin.

1. Johdanto

Informaatiotutkimus on perinteisesti jakautunut **tiedonhaun** (IR, information retrieval) ja **tiedonhankinnan** (Information seeking) tutkimukseen. Tiedonhaketutkimuksen keskeisiä tutkimusalueita ovat 1) tietokantojen sisällön ja rakenteen tutkimus, 2) tallennus- ja hakumenetelmien tutkimus, 3) käyttöliittymien tutkimus 4) käyttäjätutkimus 5) tiedonlähdetutkimus 6) evaluoinnin tutkimus 7) tietoteollisuuden tutkimus sekä 8) tiedonhaun tietotekniikoiden ja tietoliikennetekniikoiden tutkimus (Järvelin 1995, 36). Tiedonhankintatutkimuksessa on puolestaan tutkittu muun muassa tiedontarpeiden syntyä sekä kanavien ja lähteiden valintaa, käyttöä, käytön määrää ja käytön esteitä (Savolainen 1999, 74). Toisaalta informaatiotutkimus voidaan jakaa myös **järjestelmä-** ja **käyttäjakeskeiseen** suuntaukseen. Informaatiotutkimuksen kehitys liittyy läheisesti tietokoneiden ja tiedonhakujärjestelmien syntyhistoriaan, jolloin tieteenalalla on aluksi keskitytty lähinnä järjestelmäkeskeiseen tiedonhaun tutkimukseen (Saracevic 1975, 326). Tästä on eriytynyt käyttäjakeskeinen suuntaus, joka Ingwersenin (1996,12) mukaan jakautuu edelleen tiedonhaun käyttötutkimukseen ja tiedonhankintatutkimukseen.

Tutkimussuuntaukset ovat eriytyneet toisistaan riippumatta siitä, jaetaanko informaatiotutkimus tiedonhaku- ja tiedonhankintatutkimukseen tai järjestelmä- ja käyttäjakeskeiseen suuntaukseen. Suuntausten välistä yhteistyötä ovat kuitenkin ehdottaneet monet tutkijat. Kokonaisvaltaisemmassa tutkimuksessa tiedonhakua tarkasteltaisiin osana laajempaa tiedonhankinnan prosessia, jonka tarkoituksena on löytää hyödyllistä informaatiota ongelman ratkaisuun tai tehtävän suorittamiseen. (Bates 1989, 407; Hert 1996, 513; Ingwersen 1996, 42; Vakkari 1999, 820.)

Tämän tutkimuksen painopiste on lähinnä Ingwersenin (1996, 12) jaottelun mukaisessa tiedonhaun käyttötutkimuksessa. Lisäksi tiedonhakua tarkastellaan toimintona, jonka tarkoituksena on ongelman tai tehtävän ratkaisemisessa auttavan tiedon hankkiminen. Tarkastelun lähtökohtana on konstruktivistinen käsitys oppimisesta, jonka mukaan uuden tiedon oppiminen rakentuu vanhan, aiemmin omaksutun tiedon varaan. Tällöin tiedonhaku on osa oppimistapahtumaa, jossa uutta tietoa hankittaessa hyödynnetään aiemmin opittua. Tiedonhauissa voidaan hyödyntää muun muassa aiempaa aihetuntemusta, opittuja tiedonhakutaitoja sekä käytetyn tiedonhakujärjestelmän tuntemusta.

Tutkimuksen keskeisenä tarkoituksena oli selvittää, **minkä tyyppistä tietoa loppukäyttäjät odottivat tiedonhakujen avulla löytävänsä** ennalta rajatun tehtävän suorituksen eri vaiheissa sekä **minkä tyyppistä ja miten hyödyllistä tietoa loppukäyttäjät arvioivat viitteiden perusteella löy-**

täneensä. Tietotyyppjä tarkastellaan tässä työssä käyttäjakeskeisinä aiheenmukaisuuden ja relevanssin kriteereinä. Lisäksi työssä tarkastellaan tiedonhakujen onnistuneisuutta käyttäjien tekemien arvioiden pohjalta.

Tutkimukseen pyrittiin valitsemaan mahdollisimman homogeeninen ryhmä henkilöitä, joilla oli suoritettavanaan osapuilleen sama tehtävä. Tutkimusaineisto kerättiin vuoden 2000 syys- marras- kuussa. Tutkimukseen osallistui 22 Tampereen yliopiston psykologian laitoksen opiskelijaa, joilla oli tehtävänään laatia tutkimussuunnitelma suppeaa empiiristä tutkimusta varten. Kurssin aikana heidän tuli laatia psykologian aineopinnoissa ensin alustava tutkimussuunnitelma ja sen jälkeen perusteellisempi lopullinen tutkimussuunnitelma. Opiskelijoita havainnoitiin kaksi kertaa: kurssin alkuvaiheessa vähän ennen ensimmäisen alustavan tutkimussuunnitelmansa luovutusta ja kurssin lopussa viimeistellyn tutkimussuunnitelman luovutuksen aikoihin.

2. Teoreettinen tausta ja keskeiset käsitteet

Tehtävän tai ongelman ratkaisemiseen liittyy aina oppimista. Henkilö kohtaa oppimisprosessissa ongelman, joka ei ole selitettävissä tai ratkaistavissa sen hetkisen tietämyksen valossa. Oppijan on hankittava ratkaisumalleja, jotka muuttavat hänen tietämyksensä tehtävän ratkaisemisen edellyttämän mallin mukaiseksi (Engeström 1987, 46). Vaativaa tehtävää suoritettaessa joudutaan vuorostaan ratkaisemaan useita osatehtäviä ja muodostuneet uudet tietämyksen mallit saattavat herättää edelleen uusia kysymyksiä ja ongelmia. Monimutkainen ja vaativa tehtävä edellyttää siten toistuvia tiedonhakuja ja jatkuvaa tiedonhankintaa (Kuhlthau 1993, 41). Kun informaalit kanavat osoittautuvat riittämättömiksi, turvaudutaan formaaleihin tiedon hankintakeinoihin ja hakujärjestelmiin.

2.1. Aihetietämyksen muutos tiedon hankinnan ja tehtävän suorituksen eri vaiheissa

Ongelman ratkaiseminen edellyttää tehtävään liittyvien keskeisten käsitteiden ja niiden välisten suhteiden ymmärtämistä. Tehtävän suorittaminen edellyttää myös valintojen ja aiherajausten tekemistä. Kuhlthau (1991, 366; 1993, 41) kuusivaiheisen tiedonhankinnan prosessimallin (Kuvio 1) mukaan ongelman ratkaiseminen tapahtuu vaiheittain, jolloin tiedonhakijan tuntemukset, ajatukset, toiminnot ja osatehtävät muuttuvat tehtävän suorittamisen aikana.

	Tuntemukset	Ajatukset	Toiminnot	Osatehtävät
Aloitus	Epävarmuus	Yleisiä ja vaihtelevia	Taustatiedon hankkiminen	Tiedon tarpeen tunnistaminen
Valinta	Optimismi			Yleisen aiheen tunnistaminen
Tunnustelu	Hämmennys, turhautuminen, epäilykset		Relevantin tiedon hankkiminen	Aiheeseen perehtyminen
Muotoilu	Selkeytymisen tunne	Selkiytyneempiä ja tarkemmin rajattuja		Fokuksen muotoileminen
Keräys	Päämäärä-tietoisuus. luottavaisuus	Kiinnostuksen lisääntyminen	Relevantin ja rajatun tiedon hankkiminen	Tiedon koonti
Esitys	Tyytyväisyys tai pettynisyys	Selkiytyneet ja fokuoituneet		Viimeistely

Kuvio 1. Kuhlthaun (1991, 367) tiedonhankinnan prosessimallin vaiheet, ja niiden aikana ilmenevät tuntemukset, ajatukset, toiminnot ja tehtävät.

Tiedonhankinnan prosessimallin **aloitusvaiheessa** tuntemukset ovat epävarmoja, kun tehtävän suorittaja huomaa sen hetkisen tietämyksen olevan riittämätöntä kohdatun ongelman ratkaisemiseen tai tehtävän suorittamiseen. Aluksi jo tehtävän ymmärtäminenkin voi olla vaikeaa eikä tiedon tarvetta osata määritellä tai muotoilla esimerkiksi hakukysymyksiksi. Epävarmuus vaihtuu kuitenkin yleensä optimismiksi, kun löydetään sellaista ongelman aihealueeseen liittyvää tietoa, joka pystytään käsittämään aiemman tietämyksen avulla. Tiedonhankinta keskittyy aluksi lähinnä lähteiden tunnistamiseen ja löytämiseen. **Valintavaiheessa** ongelman aihealuetta pyritään kartoittamaan hankkimalla mahdollisimman paljon aihetta koskevaa yleistä taustatietoa. Koska ongelma voidaan yleensä ratkaista monella tavalla, tässä vaiheessa pohditaan myös alustavasti aiheen rajausta oman kiinnostuksen ja tehtävän asettamien resurssien pohjalta. Tällaisia resursseja voivat olla esimerkiksi tehtävän suorittamiselle asetetut vaatimukset, aikataulu sekä tarjolla olevat tiedonlähteet.

Tunnusteluvaiheessa keskeisellä sijalla on aiheeseen perehtyminen. Ongelmaa tarkastellaan edelleen yleisellä tasolla. Tietoa pyritään kuitenkin hankkimaan valikoidummin sen hyödyllisyyden perusteella. Tuntemukset ovat epäileviä ja hämmentyneitä, kunnes uusi tieto ja sen myötä esille tulleet käsitteet ja niiden väliset suhteet pystytään sisäistämään. **Muotoiluvaiheessa** hankittu tieto ja vanha aiempi käsitys aiheesta sulautuvat uudeksi opituksi tietämykseksi. Uuden tiedon omaksumisen myötä myös ajatukset ja tuntemukset selkiytyvät. Mikäli tehtävän suoritus edellyttää opitun tiedon ulkoistamista joko esittämällä tai käytäntöön soveltamalla, voidaan tässä vaiheessa ryhtyä

muotoilemaan oman esityksen **tarkempaa aiheenrajausta** (focus). Muotoiluvaiheessa kerätään vastaavasti tarkempaa aiheenrajausta hyödyttävää tietoa.

Keräilyvaiheessa herää myös syvempi kiinnostus aihetta kohtaan, kun aiheen tarkempi raja on selvillä. Tällöin voidaan päämäärätietoisesti ryhtyä kokoamaan omaa esitystä sekä hankkimaan valittua aihetta tukevaa yksityiskohtaista tietoa. Yleensä lisääntynyt kiinnostus johtaa oman esityksen laatimiseen riippumatta siitä, onko sitä tehtävän suorituksessa edellytetty. **Esitysvaiheessa** oma rajattu käsitys viimeistellään suulliseksi, kirjalliseksi tai muulla tavoin esitettäväksi työksi. Esitysvaiheessa hankitaan lähinnä epäselviksi jääneiden yksityiskohtien tarkentamisessa auttavaa tietoa.

2.2. Relevanssin käsite informaatiotutkimuksessa

Relevanssin käsitteen määrittelemine on muodostunut keskeiseksi ongelmaksi informaatiotutkimuksessa. Chamberin (1994, 34) mukaan syitä tähän ovat vaikeus tutkia ihmisten käyttäytymistä relevanssiarviointitilanteissa, relevanssin mittaamisen ongelmallisuus sekä yhtenäisen termistön puuttuminen tieteenalalla. Järjestelmäkeskeisessä tiedonhaku tutkimuksessa relevanssin käsite on perinteisesti yhdistetty aiheenmukaisuuteen (topicality) joka on objektiivisesti määritelty useiden asiantuntijoiden avulla (esim. Cranfield testit 50-luvulla ja TREC –testikokoelmat 90-luvulla). Käyttäjakeskeisen tutkimuksen alalla on puolestaan tutkittu relevanssiarviointiin vaikuttavia tekijöitä sekä esitetty relevanssin asteita ja piirteitä ilmaisevia luokitteluita loppukäyttäjien näkökulmasta (Barry 1994, 153; Barry ja Chamber 1998, 225; Bateman 1998, 26; Chamber 1991; Vakkari ja Hakala 2000, 550; Wang ja Soergel 1998, 122).

Informaatiotutkimuksessa relevanssin käsitettä on yritetty määritellä yleisesti Shannonin ja Weaverin (1963, 98) viestintämallin mukaisesti. Siinä relevanssi ilmaisee lähteen ja kohteen välisen yhteyden tehokkuutta kommunikaatioprosessissa (Saracevic 1975, 325). Järjestelmäkeskeisestä näkökulmasta tarkasteltuna lähteenä olisivat tällöin kyselyt ja kohteena informaatio-objektit (dokumentit tai niiden kuvaukset). Tällöin tehokkuutta mitataan kyselyjen ja informaatio-objektien tekstin täsmäytyksen perusteella. Useissa järjestelmäkeskeisissä tutkimuksissa mittarina on käytetty **objektiivisesti relevanteiksi**¹ todettujen dokumenttien **tarkkuutta**² ja **saantia**³. Käyttäjakeskei-

¹ Järjestelmäkeskeisessä tutkimuksessa dokumenttien relevanttius on ennalta määritetty yleensä aiheenmukaisuuden perusteella.

² Hakujärjestelmän avulla löytyneiden relevanttien dokumenttien suhde kaikkiin löytyneisiin dokumentteihin.

sessä tutkimuksessa lähteellä voidaan puolestaan tarkoittaa esimerkiksi loppukäyttäjän **tiedon tarvetta**, ja kohteella **dokumenttien tiedon tarpeen tyydyttävää informaatioisisältöä**. Loppukäyttäjän näkökulmasta tehokkuus ilmenee esimerkiksi dokumenttien tarjoaman informaation hyödyllisyytenä (utility) tai käyttäjän tyytyväisyytenä (satisfaction). (Saracevic 1975, 325 ja 328.)

Saracevic (1996, 7) on esittänyt, että tiedonhaun vuorovaikutus ilmenee **hakujärjestelmän tasolla** kyselyjen ja informaatio-objektien täsmäytyksen muodossa, **käyttäjän kognitiivisella tasolla** informaatiolähteiden⁴ tulkintoina sekä **tilannesidonnaisella tasolla** käyttäjän tehtäväkohtaisena tiedon tarpeena. Relevanssin käsite voidaan edelleen määritellä näiden vuorovaikutusten tasojen perusteella. Saracevicin (1996, 214) mukaan relevanttius voi esiintyä järjestelmärelevanssina, aiherelevanssina, kognitiivisena relevanssina, tilannesidonnaisena relevanssina ja tavoitteellisena relevanssina. **Järjestelmärelevanssi** ilmenee kyselyjen ja informaatio-objektien merkkitason vastaavuutena. **Aiherelevanssilla** tarkoitetaan kyselyn ja haettavien tekstien aiheenmukaisuuden vastaavuutta. **Kognitiivisella relevanssilla** viitataan informaatio-objektien tietosisällön sekä käyttäjän tiedon tarpeen ja tietämyksen tason vastaavuuteen kognitiivisella tasolla. **Tilannesidonnainen relevanssi** määrittyy suhteessa käyttäjän käsillä olevaan ongelmaan tai tehtävään, jolloin relevanssi mitataan ennen muuta informaation hyödyllisyyden perusteella. **Tavoitteellisella relevanssilla** tarkoitetaan käyttäjän henkilökohtaisten intressien, motivaatioiden ja pidempiaikaisten tavoitteiden vaikutusta relevanssiarvioihin.

Järjestelmäkeskeisissä tutkimuksissa on tarkasteltu ainoastaan järjestelmärelevanssia ja aiherelevanssia⁵. Tosin käyttäjäkeskeiset tutkimukset ovat osoittaneet aiheenmukaisuuden olevan myös loppukäyttäjien yleisin mainitsema valintakriteeri (Schamber 1991; Bateman 1998, 30; Vakkari ja Hakala 2000, 557; Wang ja Soergel 1998, 122). Aiheenmukaisuuden voi kuitenkin katsoa olevan myös yhtä lailla niin kognitiivisen relevanssin, tilannesidonnaisen relevanssin kuin tavoitteellisen relevanssin yksi kriteeri. Yksittäiselle käyttäjälle aiheenmukaisuus määrittyy suhteessa henkilön tietämykseen, ongelmanratkaisutilanteeseen ja laajempiin kokonaistavoitteisiin.

³ Hakujärjestelmän avulla löytyneiden relevanttien dokumenttien suhde kaikkiin tietokannassa oleviin relevantteihin dokumentteihin, kun tietokannan koko ja relevanttien dokumenttien määrä on tunnettu.

⁴ Informaatiolähteinä kognitiivisella tasolla ovat käyttäjän oma aiempi tietämys sekä dokumentit ja niiden kuvaukset (viitteet, tiivistelmät, asiassanat jne.).

⁵ Järjestelmäkeskeisessä tutkimuksessa aiherelevanssi määritellään yleensä objektiivisesti (usean asiantuntijan mukaan) hyväksytyyn aiheenmukaisuuden perusteella.

Tässä tutkimuksessa keskitytään tarkastelemaan relevanssia lähinnä tilannesidonnaisesta näkökulmasta. Siksi haastatelluilta tiedusteltiin viitteiden hyödyllisyyttä suhteessa heidän tehtäväänsä. Tutkimuksessa on käytetty termejä **hyödyllisyys** (utility) ja **anti** (contribution) tarkoitettaessa käyttäjien tilannesidonnaista relevanssia.

2.3. Käyttäjakeskeisen relevanssin eri asteet

Relevanssia käsittelevissä tutkimuksissa viitteet ja dokumentit on tavallisesti luokiteltu joko relevanteiksi tai epärelevanteiksi (Spink, Greisdorf ja Bateman 1998, 603). Saracevicin (1996, 216) mukaan informaatio-objektien relevanttius ilmenee käyttäjälle eri asteisena joko täsmätyksen, aiheenmukaisuuden, tietämyksen, tilannesidonnaisuuden tai tavoitteellisuuden tasolla. Tällöin informaatio-objektit voivat olla käyttäjälle joko täysin relevantteja, kun kaikki relevanssin eri tasojen ehdot täyttyvät, tai täysin epärelevantteja, kun mikään näistä ehdoista ei täyty. Osittainen relevanttius (tai epärelevanttius) puolestaan ilmenee esimerkiksi silloin, kun informaatio-objekti käyttäjän mielestä vastaa kyselyä, on aiheenmukainen ja relevantti myös kognitiivisella tasolla, mutta ei täytä tilanteen tai henkilökohtaisten mieltymysten asettamia vaatimuksia (vrt. Spink et al. 1998, 604).

On kuitenkin huomattava, että tiedonhakutilanteessa käyttäjä voi antaa dokumenttien tai niiden kuvausten perusteella ainoastaan **heikkoja relevanssiarvioita** (Harter 1992, 607). Vasta tutustuttuaan paremmin dokumenttien informaatio-sisältöön käyttäjä pystyy esittämään käsityksensä dokumenttien todellisesta hyödyllisyydestä. Dokumenttien hyödyllisyys voi siten myöhemmin muuttua suhteessa tilanteeseen ja kokonaistavoitteisiin. Tiedonhakupöydällä käyttäjä voi esimerkiksi valita dokumentin pelkästään aiheenmukaisuuden perusteella relevantiksi. Tilannesidonnainen ja henkilökohtaisista mieltymyksistä riippuva relevanttius pystytään usein päättelemään kunnolla vasta dokumentin informaatio-sisältöön perehtymällä. Dokumentit voivat olla osittain relevantteja käyttäjälle myös tiettyjen osien ja ominaisuuksien perusteella. Dokumenttien yksittäiset osiot tai piirteet voivat tarjota relevanttia tietoa tai auttaa tiedonhankinnassa eteenpäin.

2.4. Informaatio-objektien valintaan vaikuttavat kriteerit

Useiden tutkijoiden tavoitteena on ollut kartoittaa ja luokitella informaatio-objektien valintaan vaikuttavia kriteereitä (Schamber 1991; Park 1993, 330; Barry 1994, 149; Bateman 1998, 126; Vakkari ja Hakala 2000, 557; Wang ja Soergel 1998, 122). Tutkimukset tosin vahvistavat aiheenmukaisuuden olevan kaikkein yleisin käyttäjien mainitsema valintakriteeri, mutta myös muut kriteerit vaikuttavat arvioihin informaatio-objektien hyödyllisyydestä (Wang ja Soergel 1998,129). Va-

lintakriteereitä ovat muun muassa arviot dokumenttien **hankittavuudesta, fyysisistä ominaisuuksista, lähteitten ominaisuuksista** sekä **informaatiosisällön ominaisuuksista** (Barry 1994, 153). Valintaan vaikuttavat myös käyttäjän **kokemus, havainnot** ja **tietämyksen taso** (Park 1993, 329). Kokemukset ja tietämys voivat liittyä esimerkiksi ammatti- tai tieteenalaan, aiempiin tehtäviin tai koulutukseen. Havainnot ovat läheisesti yhteydessä tietämykseen ja kokemukseen. Käyttäjä tekee valintoja esimerkiksi saatavuuden tai tuoreuden perusteella, koska hänellä on aiempaa kokemusta näiden tekijöiden vaikutuksesta. **Aihekohtainen tietämys** voi sen sijaan vaihdella riippuen **käyttötilanteesta, tehtävän vaiheesta** ja **henkilökohtaisista mieltymyksistä**.

Tutkijat ovat luokitelleet valintakriteerit lähinnä tutkimusten tavoitteiden perusteella, mutta luokitusten taustalla vaikuttaa myös tutkijoiden määritelmät aiheenmukaisuudesta. Muun muassa Barry ja Schamber (1998, 222) eivät sisällytä aiheenmukaisuutta suoranaisesti lainkaan kriteeriluokituksiinsa, sillä heidän tavoitteenaan on ollut kartoittaa, minkä muiden kriteereiden paitsi aiheenmukaisuuden perusteella käyttäjät tekevät relevanssiarvioita. Wang ja Soergel (1998, 122) puolestaan erottelevat kriteeriluokituksessaan aiheenmukaisuuden selkeästi rinnasteisena muille kriteereille. Tutkimuksessaan he rinnastavat aiheenmukaisuuden Barryn (1994, 153) informaatisällön ja Schamberin (1991, 129) spesifiyden valintakriteereiden kanssa (Wang ja Soergel 1998, 130).

Wangin ja Soergelin (1998, 128) mukaan käyttäjien on myös todettu tekevän toisinaan valintoja näennäisesti täysin riippumatta aiheenmukaisuudesta, mutta aiheenmukaisuus on yleensä aina ehtona myös muille kriteereille. Tutkimuksessa kartoitettiin valintakriteereitä ainoastaan käyttäjien tekemien mainintojen perusteella, joten kaikki yksittäisen informaatio-objektin valintaan vaikuttaneet kriteerit eivät välttämättä tulleet esille haastattelutilanteessa (mt. 119). Barry ja Schamber (1998, 222) eivät puolestaan ottaneet lainkaan kantaa aiheenmukaisuuden osuuteen käyttäjien tehdessä valintoja.

2.5. Haetun tiedon anti hyödyllisyyden mittana

Wangin ja Soergelin (1998, 121) mukaan viitetietokantahauissa informaatio-objektien hyödyllisyys päätellään dokumenttien oletetun arvon (value) perusteella. Tutkijoiden mukaan dokumenttien hyödyllisyyttä voidaan esimerkiksi arvioida informaation **tiedollisen arvon** (epistemic value), **tiedon käyttökelpoisuuden** (functional value), **käyttökelpoisuuden rajoitusten** (conditional value), **sosiaalisen tilanteen** (social value) ja **tunneperäisten tekijöiden** (emotional value) perusteella. Informaation tiedollisella arvolla tarkoitetaan oletusta siitä, miten dokumenttien avulla pystytään tyy-

dyttämään tiedon tarve. Tiedon käyttökelpoisuudella tarkoitetaan oletusta dokumenttien annista (contribution) omalle työlle. Käyttökelpoisuuden ehdoiksi Wang ja Soergel mainitsevat lähinnä tilanteet, joissa dokumenttien hyöty on vielä avoin. Sosiaalisilla tilanteilla tarkoitetaan puolestaan hyödyllisyyden riippuvuutta sosiaalisista suhteista tai työympäristöistä. Tunneperäisillä tekijöillä tutkijat tarkoittavat käyttäjän omien arvojen ja asenteiden vaikutusta relevanssiarvioihin.

Wangin ja Soergelin (1998, 123) tutkimustulosten perusteella käyttäjien yleisimmäksi valintakriteeriksi mainitsema aiheenmukaisuus oli sidoksissa informaation tiedolliseen arvoon, käyttökelpoisuuteen sekä käyttökelpoisuuden rajoituksiin. Toiseksi yleisimmin mainittu valintakriteeri oli arvio dokumenttien informaation tiedollisesta suuntautuneisuudesta (orientation) tai tasosta (level), jonka käyttäjät puolestaan arvioivat olevan sidoksissa tiedon käyttökelpoisuuteen ja sen rajoituksiin. Lisäksi käyttäjät tekivät päätelmiä dokumenttien suuntautuneisuudesta yleensä vasta todettuaan ne myös aiheenmukaisiksi. Vaikka viitteiden ja dokumenttien valintaan voivat siis vaikuttaa monet kriteerit, käyttäjä arvioi dokumenttien hyötyä lähinnä tiedollisen arvon ja käyttökelpoisuuden perusteella.

Vakkarin (2000, 2) mukaan tiedonhaku- ja hankintahetkellä dokumenttien hyödyllisyyttä arvioidaan tietotyyppien perusteella. Käyttäjä hakee siis erityisesti sen tyyppistä tietoa, jonka hän arvelee auttavan käsiteltävän ongelman tai tehtävän ratkaisussa. Lisäksi tietotyyppien hyöty on sidoksissa tehtävän vaiheeseen. Tehtävän eri vaiheissa esiin tulevien ongelmien ratkaisussa tarvitaan siten eri tyyppistä tietoa.

Tietotyypit voidaan jaotella ensinnäkin tiedon spesifiyden perusteella. Tehtävän alussa aihetietämyksen ollessa vähäistä haetaan yleistä taustatietoa, jota ei pystytä kuvailemaan yksityiskohtaisesti ja jota käytetään tehtävän rajaamisessa (Kuhlthau 1993, 58; Vakkari 2000, 5; Su 1994, 214). Aiheetietämyksen lisääntyessä haetaan puolestaan yksityiskohtaisempaa tietoa, joka käsittelee tehtävän kannalta keskeisiä muuttujia (Kuhlthau 1993, 58; Vakkari 2000, 5). Tiedon spesifiys esiintyy myös Chamberlinin (1991, 129) tekemässä valintakriteerien luokituksessa. Tällöin tarkoitetaan käyttäjän tarpeen kannalta riittävän yksityiskohtaista tietoa.

Hakala (2000, 19) ja Vakkari (2000, 5) nimeävät lisäksi kolme muuta käyttäjän relevanssivalintoihin vaikuttavaa tiedon tyyppiä: teoritieto, empiriatieto ja metoditieto. Jaottelu on peräisin Raution ja Suhosen (1981, 25) käyttämästä väitöskirjojen viittausten ja lähteiden tietotyyppien mukaisesta luokittelusta. **Teoriatiedolla** tarkoitetaan tutkimuskohdetta tai tieteenalaa koskevia keskeisiä teori-

oita, hypoteeseja, yleistyksiä ja käsitteitä, jotka koskevat tutkimuskohdetta tai jotakin laajempaa todellisuuden osaa. **Empiriatiedolla** tarkoitetaan tietoa, joka ei ole saavuttanut teoreettisen yleistyksen tasoa. **Metoditiedolla** tarkoitetaan tietoa tutkimuksessa käytettävistä menetelmistä, tekniikoista ja käytännöistä. Vakkari (2000, 6) on osoittanut käyttäjien tekemien viitteiden tietotyyppiarvioiden perusteella, että teorian tietoa tarvitaan lähinnä tehtävän alussa ja keskivaiheilla, kun taas empiria- ja menetelmätiedon tarve lisääntyy tehtävän loppupuolella.

Myös Wangin ja Soergelin (1998, 123) tutkimuksessa käyttäjät erittelivät tiedonhaun yhteydessä dokumenttien informaatioisällön suuntautuneisuutta muun muassa teorian tiedon, empiriatiedon ja menetelmätiedon perusteella. Tutkimustulosten perusteella tiedon suuntautuneisuutta tarkastellaan suhteessa käyttötarkoitukseen ja tilanteen asettamiin käyttötarkoituksen ehtoihin. Näin ollen myös Wangin ja Soergelin tutkimuksen perusteella voidaan päätellä, että teoria-, empiria ja metoditieto ovat sidoksissa tehtävän vaiheeseen.

2.6. Tiedonhaun yleinen onnistuneisuus käyttäjän näkökulmasta

Yleensä informaatiotutkimuksessa haun onnistuneisuutta arvioidaan järjestelmäkeskeisen relevanssikäsityksen mukaisesti **tarkkuuden** ja **saannin** avulla. Käyttäjakeskeisestä näkökulmasta tarkasteltuna tulisi kuitenkin käyttää mittareita, jotka ilmaisevat paremmin tiedon loppukäyttäjän käsityksiä haun onnistuneisuudesta. Sun (1994, 209) mukaan käyttäjät eivät pidä tärkeänä määrällistä tarkkuutta. Sen sijaan hakijan tyytyväisyys tarkkuuteen mittaa paremmin haun onnistuneisuutta.

Su arvioi käyttäjien esittämien kommenttien perusteella laadun olevan ylipäätään määrää tärkeämpi tekijä. Tämä ei tarkoita sitä, etteikö käyttäjä olisi sitä tyytyväisempi mitä enemmän hyödyllisiä viitteitä hän löytää. Käyttäjät ovat ennemminkin huolestuneita saannista. Käyttäjälähtöisessä tutkimuksessa on kuitenkin mahdoton mitata saantia suoraan luettamatta kaikkia dokumentteja käyttäjällä. Sen sijaan Su mittasi käyttäjien omaa käsitystä saannista kysymällä heiltä, kuinka tyytyväisiä ja luottavaisia he olivat haun **kattavuuteen** (completeness⁶) sekä miten tärkeäksi he kattavuuden kokivat. Erityisesti haettaessa tietoa vaativaa tieteellistä tehtävää varten käyttäjät pyrkivät varmistamaan, ettei oleellisia dokumentteja jäisi löytymättä. (Su 1994, 211.)

⁶ Käännös omani, sillä mielestäni suorat käännökset Su (1994, 213) käyttämälle termille completeness (täydellisyys, kokonaisuus, eheys, ehjyys) eivät vastaa asiayhteyttä.

Ne jotka tekivät tiedonhakuja joko tutkielmatyötä tai apurahahakemusta varten, korostivat erityisesti saannin merkitystä. Heidän tavoitteenaan oli varmistaa tutkimuksensa uutuus sekä töidensä lähdeaineiston aukottomuus. Kuitenkin myös monet yksinkertaisempia tehtäviä suorittaneet pitivät kattavuutta tärkeänä eri syistä. Keskeistä kaikille käyttäjille oli se, että he työskentelivät akateemisessa työympäristössä, jossa arvostetaan lähteiden tarkkuutta. Myös tehtävän vaiheella on merkitystä haun onnistuneisuuden arvioihin. Tehtävän alkuvaiheessa oleville hakujen kattavuus ja vastaavasti saanti näytti olevan vähemmän tärkeää. Käyttäjät ilmaisivat tällöin hakevansa alustavaa tai yleistä tietoa tarkastelemastaan aiheesta. (Su 1994, 214.)

3. Aikaisempi tutkimus

Muun muassa Spink, Greisdorf ja Bateman (1998, 599) ovat tutkineet aihetietämyksen vaikutusta käyttäjien arvioihin relevanssin asteesta. Kuhlthau (1991, 366) on puolestaan tutkinut aihetietämyksen muutosta ja tiedonhankinnan vaiheita. Hänen laatimansa tiedonhankinnan prosessimalli on luonut pohjaa myöhemmille tiedonhaun- ja tiedonhankinnan pitkittäistutkimuksille. Kuhlthau sivuaa lisäksi aihetietämyksen vaikutusta haetun tiedon spesifiteetin. Bateman (1998, 23) on tutkinut valintakriteereiden tärkeyttä käyttäjille tiedonhankinnan eri vaiheissa. Hakala ja Vakkari (Hakala 1999, 33; Vakkari 2000, 4; Vakkari ja Hakala 2000, 547) ovat puolestaan tutkineet viitteiden tietotyypin- ja relevanssiarvioiden muutosta suhteessa aihetietämyksen kasvuun ja tiedon hankinnan vaiheeseen.

3.1. Relevanssin asteet suhteessa aihetietämyksen määrään

Spink, Greisdorf ja Bateman (1998, 607) vertailevat tutkimuksessaan relevanssin eri asteiden suhdetta aihetietämyksen määrään ja aihetietämyksen kasvuun kolmessa käyttäjäryhmässä. Ryhmät koostuivat pääasiassa yliopisto-opiskelijoista, jotka tekivät ensimmäisiä tiedonhakuja omaa työtään varten (mt. 609). Kahdessa ryhmässä oli ainoastaan yliopisto-opiskelijoita, jotka tekivät hakunsa itse. Kolmannessa ryhmässä oli opiskelijoiden lisäksi yliopiston henkilökuntaa. Tässä ryhmässä kokeneet välittäjät tekivät haut käyttäjien kuvausten perusteella, käyttäjien ollessa kuitenkin itse läsnä tiedonhakutilanteessa. Käyttäjää pyydettiin arvioimaan löytyneiden informaatio-objektien hyödyllisyyttä asteikolla hyödyllinen, osittain hyödyllinen ja ei lainkaan hyödyllinen. Lisäksi kysyttiin käyttäjien käsitystä oman aihetietämyksensä määrästä esillä olevan ongelman suhteen, sekä missä määrin he havaitsivat aihetietämyksen lisääntyneen haun seurauksena (Spink et al. 1998,

607). Käyttäjien käsitystä aihetietämyksen määrästä ja muutoksesta verrattiin viitteiden määriin hyödyllisyysasteittain.

Tutkijat eivät varsinaisesti perustele, miksi mainittuja käyttäjäryhmiä verrataan keskenään. Tulokset kuitenkin osoittavat, että tiedonhauissa välittäjiä apunaan käyttäneen ryhmän ne käyttäjät, jotka arvioivat tietävänsä omasta aiheestaan paljon, valitsivat vastaavasti paljon täysin relevantteja viitteitä. Toisessa ryhmässä puolestaan paljon osittain relevantteja viitteitä löytäneet käyttäjät arvioivat aihetietämyksen lisääntyneen haun seurauksena⁷. Tutkijat arvioivat osittain relevanttien dokumenttien voivan tarjota täysin relevantteja dokumentteja todennäköisemmin käsitelmien muuttamiseen johtavaa uutta tietoa. Tutkijoiden mukaan erityisesti vähän omasta aiheestaan tietävät käyttäjät valitsisivat tällaisia käsitelmien muuttamisessa auttavia osittain relevantteja informaatio-objekteja. Vastaavasti omasta aiheestaan paljon tietävät, valitsisivat etupäässä tutkimusongelman tarkentamisessa auttavia täysin relevantteja viitteitä.

Tutkimuksessa ei tarkasteltu hakukokemuksen vaikutusta tuloksiin. Ryhmä, jossa esiintyi voimakasta korrelaatiota aihetietämyksen ja täysin relevanttien informaatio-objektien välillä, käytti haussa apunaan kokeneita välittäjiä. (Spink et al. 1998, 610). Hsieh-Yeen (1993, 168) mukaan hakukokemuksella on selvä myötävaikutus hakujen muotoilussa - oli aihe sitten outo tai tuttu. Kokeneet hakijat yhdistävät tehokkaammin erityisesti synonyymeja hakulausekkeisiinsa sekä hyödyntävät mm. tesaarusta. Kokemattomat hakijat eivät edes tunne tarvetta synonyymien yhdistämiseen. Hsieh-Yee ei tarkastele artikkelissaan hakukokemuksen vaikutusta hakutuloksiin, mutta oletettavasti hakukokemus saattaa auttaa löytämään tehtävän joka vaiheessa enemmän täysin hyödyllisiä viitteitä.

Spink, Greisdorf ja Bateman (1998) eivät suoraan tarkenna, tarkoittavatko he aihetietämyksellä käyttäjien alakohtaista aihetietämystä (ammatti- tai tieteenala) vai tehtäväkohtaista aihetietämystä. Tutkimuksesta ei käy ilmi, kuinka suuri osuus käyttäjistä oli henkilökuntaan kuuluvia ja poikkesiko heidän arvionsa aihetietämyksestä olennaisesti opiskelijoiden arvioista. Lisäksi tutkimuksessa ei erotella myöskään yliopisto-opiskelijoita esimerkiksi sen perusteella, opiskelevatko he perusteita, aineopintoja vai syventäviä. Tutkimustulosten perusteella aihetietämyksen ja täysin relevanttien informaatio-objektien välillä oli kuitenkin merkitsevä korrelaatio ryhmässä, jossa osa käyttäjistä oli

⁷ Kysely tehtiin tiedonhakujen jälkeen, ja käyttäjiltä kysyttiin tietämyksen muutosta haun aikaisen vuorovaikutuksen ja/tai palauteprosessin seurauksena (mt. 607). Artikkelista ei käy ilmi, missä määrin aikaa haastatellut saivat käyttää informaatio-objektien tarkasteluun ja oliko kyseessä viitteet vai dokumentit.

henkilökuntaa. Kun tutkimukseen valittiin kuitenkin ainoastaan sellaisia käyttäjiä, jotka tekivät ensimmäisiä tiedonhakuja omasta aiheestaan, voidaan aihetietämyksen olettaa mahdollisesti kuvastaneen lähinnä heidän käsitystään omasta alakohtaisesta perehtyneisyydestä.

3.2. Tiedonhankinnan prosessimalli pitkittäistutkimuksen perustana

Kuhlthau (1991 ja 1993) keskittyy tutkimuksissaan ensisijaisesti käyttäjien tiedonhankinnan vaiheiden mallintamiseen tehtävän suorituksen aikana. Kuhlthau sivuaa kuitenkin myös aihetietämyksen vaikutusta haetun tiedon tyyppiin (Kuhlthau 1993, 58). Tiedonhankinnan prosessimallin perusteella käyttäjän tiedonhankinnan vaihe vaikuttaa siihen, miten laaja-alaista tai yksityiskohtaista tietoa käyttäjä kulloinkin hakee ja hankkii. Keskeistä tutkimukselle kuitenkin on, että se on luonut pohjaa ja tarjonnut välineitä tiedonhaun- ja hankinnan pitkittäiselle tutkimukselle. Kuhlthauin mallia tiedonhankinnan vaiheista ovat edelleen hyödyntäneet tutkimuksissaan Bateman (1998, 23) sekä Hakala ja Vakkari (Hakala 1999, 8; Vakkari 2000, 4; Vakkari ja Hakala 2000, 542).

3.3. Aihetietämyksen lisääntyminen tiedonhankinnan eri vaiheissa

Batemanin (1998, 26) tutkimuksessa käyttäjiä pyydettiin itse arvioimaan, missä vaiheessa Kuhlthauin tiedonhankinnan prosessimallia he olivat sekä miten tärkeänä he pitivät eri valintakriteereitä tiedonhankinnan eri vaiheissa. Tutkimuksessa käytettiin pohjana sekä Schamberin (1991) että Barryn (1994, 153) valintakriteerien luokituksia⁸. Arvioinnit tehtiin kolmessa vaiheessa: viitteiden perusteella tiedonhaun yhteydessä, kokotekstitiedostojen silmäilyn perusteella dokumenttien hankinnan yhteydessä ja dokumenttien informaattiosisällön perusteella, kun dokumentteihin oli perehdytty syvällisemmin. Tutkimusasetelma ei täysin vastaa Kuhlthauin tutkimusta, jossa henkilöt hankkivat tietoa pidemmällä aikavälillä ja tekivät toistuvia tiedonhakuja. Batemanin asetelman voi silti ajatella vastaavan pienimuotoisen tehtävän suoritusta. Siinä käyttäjä ongelman kohdattuaan tekee ensimmäisen relevanssiarvion tiedonhaun yhteydessä, sitten kokotekstien perusteella informaatioobjektien hankinnan (obtaining) yhteydessä ja lopuksi ratkaisee ongelman perehtymällä dokumenttien informaattiosisällön perusteellisemmin.

Kuhlthauin (1993) mallin osalta Batemanin (1998, 28) tutkimustulokset osoittavat, että käyttäjä ei pysty muodostamaan esitystä informaatioobjektien perusteella perehtymättä ensin niiden sisältöön.

Käyttäjät arvioivat olevansa esitysvaiheessa vasta luettuaan dokumentteja. Lisäksi tutkimuksen mukaan kaikki käyttäjät arvioivat dokumentteja luettuaan olevansa vähintään muotoiluvaiheessa, jolloin esimerkiksi tarkasteltavan aiheen valinta on jo täysin selvillä. Sen sijaan arviot valintakriteerien tärkeydestä eivät juuri muuttuneet tiedonhankinnan aikana. Aiheenmukaisuus arvioitiin tärkeimmäksi valintakriteeriksi kaikissa mitatuissa tiedon hankinnan vaiheissa, mutta myös informaation laatua (hyvin kirjoitettu, luotettava, tarkka, ymmärrettävä, johdonmukainen) pidettiin tärkeinä (Bateman 1998, 30).

3.4. Tiedon tarve vaativan tehtävän suorituksen eri vaiheissa

Hakalan ja Vakkarin (Hakala 1999; Vakkari 2000; Vakkari ja Hakala 2000) tutkimuksen keskeisiä tavoitteita oli selvittää, miten hakijoiden arviot viitteiden hyödyllisyydestä, aiheenmukaisuudesta ja valintakriteerien osuvuudesta muuttuvat tehtävän suorituksen aikana (Hakala 1999, 33). Tutkimuksessa haastateltiin 11:tä opintojen loppupuolella olevaa Tampereen yliopiston informaatiotutkimuksen pääaineopiskelijaa, joilla oli tehtävänä tutkimussuunnitelman laatiminen pro gradu –työtä varten.

Käyttäjien tekemiä tiedonhakuja seurattiin tehtävän kolmessa eri vaiheessa: tehtävän alussa, sen keskivaiheilla ja lopussa. Käyttäjien alakohtainen tietämys oli pyritty vakioimaan valitsemalla opiskelijoita jotka ovat samalta tieteenalalta, tekevät samantyyppistä tehtävää ja ovat opinnoissaan joko-kin samassa vaiheessa. Tiedonhauissa he käyttivät Dialogin LISA-tietokantaa ja hakujen tietokoneen näytöllä näkyvät vaiheet sekä hakijoiden ääneen ajattelu tallennettiin videonauhalle. Ennen hakuja ja niiden jälkeen opiskelijoita haastateltiin ja tehtäväkohtaisen aihetietämyksen määrää mitattiin jokaisella hakukerralla Kuhlthau (1993, 96) Process Survey –lomakkeen avulla. Lomakkeella ei suoranaisesti mitata tiedon määrää, vaan sen avulla voidaan muun muassa varmistua käyttäjän tietämyksen tasosta (yleistä vs. yksityiskohtaista) ja tietämyksen jäsentyneisyydestä. Aihetietämyksen ollessa yleistä ja jäsentymätöntä, tiedetään aiheesta luonnollisesti vähemmän, kuin jos aihetietämys on yksityiskohtaista ja aiheen kannalta keskeiset käsitteet ovat tuttuja. Dokumenttien hyödyllisyyttä pyydettiin arvioimaan löytyneiden viitteiden perusteella käyttäen asteikkoa täysin relevantti, osittain relevantti ja ei lainkaan relevantti. Lisäksi opiskelijat pitivät hakupäiväkirjaa, ja nämä sekä opiskelijoiden valmiit tutkimussuunnitelmat kerättiin tehtävän lopussa.

⁸ Alunperin Bateman viittaa Chamberin 1991 ja Barryn 1993 tohtorinväitöskirjoihin, mutta oletan Chamberin ja Barryn käyttäneen samoja luokituksia myös viittaamissani lähteissä.

Tulosten perusteella käyttäjät ensinnäkin etenivät tehtävän suorituksessa melko johdonmukaisesti Kuhlthaun mallin mukaisesti (Hakala 1999, 50). Ensimmäisellä hakukerralla kaikki käyttäjät olivat tosin valinneet jo aiheen ja puolet käyttäjistä olivat ehtineet perehtyä alustavasti aihepiiriin. Yksikään käyttäjä ei ollut kuitenkaan vielä ehtinyt muodostaa yksityiskohtaisempaa jäsenystä valitsemastaan aiheesta. Toisella hakukerralla kaikki olivat perehtyneet aiheeseensa ja useat olivat ehtineet muodostaa ainakin jonkinasteisen jäsenyyksen omalle esitykselleen. Tutkimuksen viimeinen hakukerta osui vaiheeseen, jossa osa oli jo ehtinyt saada oman esityksensä valmiiksi. Ne joilla tehtävä oli vielä työn alla, keräsivät lähinnä yksityiskohtaisempaa tietoa työnsä viimeistelemiseksi.

Hyödyllisiä viitteitä käyttäjät löysivät eniten tehtävän alussa. Kun mitattiin viitteiden hyödyllisyyttä asteikolla hyödyllinen ja osittain hyödyllinen, käyttäjät löysivät alussa selvästi enemmän täysin hyödyllisiksi arvioituja viitteitä. Osittain hyödyllisten viitteiden osuus kasvoi tehtävän aikana, ja tehtävän lopussa enemmän viitteitä arvioitiin osittain hyödyllisiksi kuin täysin hyödyllisiksi. Vakkarin (2000, 4) mukaan tulokset tukevat Kuhlthaun (1993) mallia myös siltä osin, että tehtävän alussa haettiin lähinnä yleistä taustatietoa ja teoretietoa. Tehtävän lopulla puolestaan korostui tutkimusmenetelmien, empiirisen tiedon ja yksityiskohtaisen tiedon merkitys. Lisäksi aiheenmukaisuuden todettiin olevan käytetyin kriteeri viitteiden hyödyllisyyttä arvioitaessa.

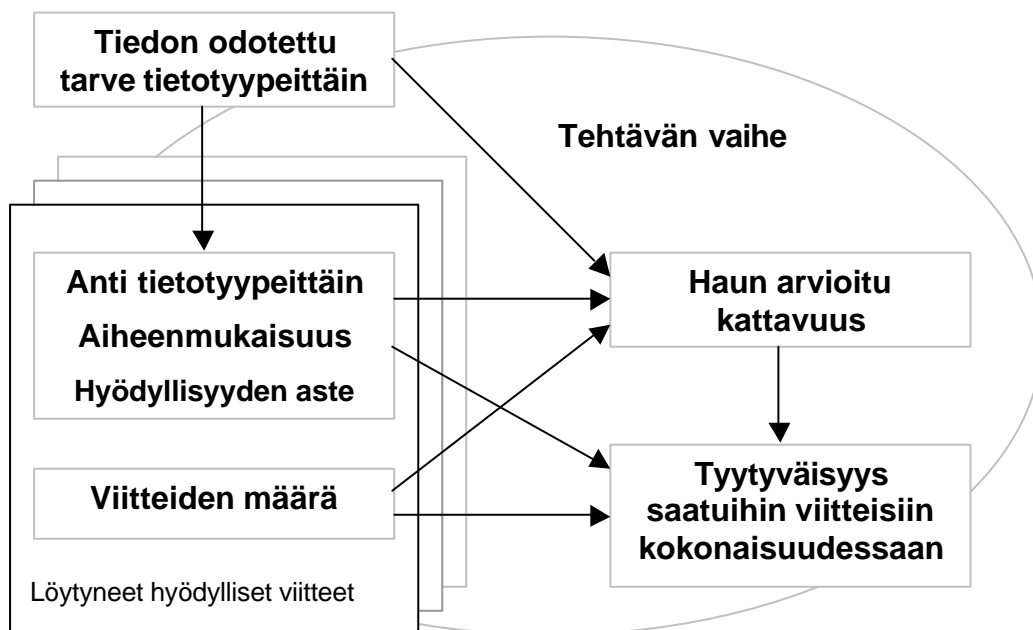
Tehtävän alussa, kun aihetietämys oli vähäistä, arvioitiin enemmän viitteistä täysin hyödyllisiksi kuin osittain hyödyllisiksi (Vakkari ja Hakala 2000, 549). Tehtävän keskivaiheilla ja lopussa hyväksyttiin kaiken kaikkiaan vähemmän viitteitä hyödyllisiksi kuin alussa, ja osittain relevanttien viitteiden osuus kasvoi loppua kohden. Siis mitä enemmän aiheesta tiedetään sitä vähemmän viitteitä hyväksytään hyödyllisiksi ja erityisesti täysin hyödyllisiksi. Vakkarin ja Hakalan (2000, 550) mukaan tämä johtuu aihetietämyksen lisääntymisestä, jolloin vastaavasti myös käyttäjien kyky tehdä eroja relevanssin asteen suhteen lisääntyy ja käyttäjät tulevat samalla myös kriittisemmiksi viitteitä valitessaan.

Vakkarin ja Hakalan (2000, 549) tutkimustulokset poikkeavat Spinkin, Greisdorfin ja Batemanin (1998, 610) tekemästä havainnosta, jonka mukaan aihetietämyksen lisääntyessä hyväksyttäisiin myös enemmän informaatio-objekteja täysin hyödylliseksi. Näitä kahta tutkimusta vertailtaessa tulee kuitenkin kiinnittää erityistä huomiota aihetietämyksen määrittelyyn. Vakkari ja Hakala (2000, 544) rajaavat varsin selkeästi aihetietämyksen nimenomaan tehtäväkeskeiseksi, kun taas Spink,

Greisdorf ja Bateman (1998, 607) jättävät aihetietämyksen melko lailla määrittelemättä, vaikka puhuvatkin aihetietämyksestä esillä olevan tehtävän suhteen.

4. Tutkimuskysymykset ja hypoteesit

Tutkimuksen keskeisenä tarkoituksena oli selvittää, **minkä tyyppistä tietoa käyttäjät odottivat löytävänsä ja minkä tyyppistä tietoa he arvioivat viitteiden perusteella löytäneensä tehtävän eri vaiheissa.** Lisäksi selvitettiin **miten tehtävän suorituksen vaihe on yhteydessä käyttäjien arvioihin viitteiden hyödyllisyydestä, aiheenmukaisuudesta ja tiedonhakujen onnistuneisuudesta yleensä.** Myös eri muuttujien keskinäisiä suhteita tarkasteltiin vertaamalla esimerkiksi tietotyyppien ennako-odotuksia arvioihin löytyneiden viitteiden tietotyyppien annista tehtävän eri vaiheissa. Kuviossa 2 on esitetty tutkimuksessa tarkastellut keskeiset käsitteet ja niiden väliset suhteet.



Kuvio 2. Tutkimuksessa tarkasteltujen muuttujien väliset suhteet.

Kuvion 2 mukaisesti käyttäjät odottavat löytävänsä tiedonhaun avulla tietyn tyyppistä oman työn ja tehtävän vaiheen kannalta aiheenmukaista ja hyödyllistä tietoa. Tiedon tarvetta mitattiin tutkimuksessa tietotyyppien avulla, koska käyttäjien voi olettaa hakevan eri tyyppistä tietoa työn eri vaiheissa. Tutkimuksessa oletettiin, että käyttäjät kuitenkin kaipaavat mahdollisimman aiheenmukaista ja hyödyllistä tietoa, joten käyttäjien odotuksia aiheenmukaisuuden ja hyödyllisyyden suhteen ei mitattu. Lisäksi oletettiin, että tiedonhaun aikana käyttäjät valitsevat viitteitä, joiden he arvioivat

auttavan informaatio-sisällöltään tiedon odotettua tarvetta vastaavien dokumenttien hankkimisessa. Arviot viitteiden tietotyyppiannista, aiheenmukaisuudesta ja hyödyllisyydestä ovat siis samalla arvioita dokumenttien oletetusta hyödyllisyydestä. Haun onnistuneisuutta kokonaisuudessaan oletettiin arvioitavan tiedon odotetun tarpeen, löytyneiden viitteiden annin ja löytyneiden viitteiden määrän perusteella. Käyttäjien oletettiin olevan tyytyväisiä saamiinsa viitteisiin kokonaisuudessaan sekä kokevan haut kattaviksi, jos he löytävät paljon aiheenmukaisia, hyödyllisiä ja odotetun tyyppistä tietoa tarjoavia viitteitä. Kokemuksen kattavuudesta oletettiin lisäävän edelleen tyytyväisyyttä saatuihin viitteisiin kokonaisuudessaan. Käyttäjien oletettiin siis olevan saamiinsa viitteisiin sitä tyytyväisempiä mitä kattavammaksi he haun kokevat.

4.1. Tiedon odotettu tarve arvioituna tietotyypeittäin

Tutkimuksessa selvitettiin, minkä tyyppistä tietoa psykologian opiskelijat odottivat tarvitsevansa tehtävän eri vaiheissa ja minkä tyyppistä tietoa he arvioivat löytyneiden viitteiden antavan työlleen. Lisäksi haluttiin vertailla, miten löytyneiden viitteiden tietotyyppiarviot erosivat ennen hakua esitetyistä tietotyyppien odotetusta vastaavuudesta.

Käyttäjakeskeisten tutkimusten perusteella aiheenmukaisuus on loppukäyttäjien useimmin mainitsema informaatio-objektien valintaperuste (Bateman 1998, 30; Schamber 1991; Vakkari ja Hakala 2000, 557; Wang ja Soergel 1998, 122). Tässä tutkimuksessa oletettiin, että aiheenmukaisuus on yhteydessä tiedon tarpeeseen, jolloin haetaan tietyn tyyppistä tietoa jonkin ongelman tai tehtävän ratkaisemiseen. Lisäksi oletettiin, että paljon tiedonhankintaa ja useita tiedonhakuja vaativan tehtävän eri vaiheissa tarvitaan vastaavasti monen tyyppistä tietoa riippuen ongelman jäsentyneisyydestä.

Tietotyypit voidaan jaotella esimerkiksi tiedon **yksityiskohtaisuuden** ja **käyttötarkoituksen** perusteella. Tehtävän alussa haetaan pääasiassa yleistä, tehtävän rajauksessa auttavaa taustatietoa, jota ei pystytä yksityiskohtaisesti kuvailemaan. Tehtävän lopussa kasvaa yksityiskohtaisemman, tehtävän keskeisiä muuttujia käsittelevän tiedon tarve. Tiedon yksityiskohtaisuus on siten sidoksissa aihetietämykseen ja tehtävän vaiheeseen. (Kuhlthau 1993, 365; Vakkari 2000, 5.)

Kuhlthau (1993) prosessimallin perusteella laadittu tiedon yksityiskohtaisuuden mukainen tietotyypijaottelu:

Yleinen taustatieto - tietoa jota ei pystytä kuvailemaan yksityiskohtaisesti ja jota käytetään tehtävän rajaamisessa.

Yksityiskohtainen tieto - tietoa joka on esitetty yksityiskohtaisesti ja käsittelee tehtävän kannalta keskeisiä muuttujia.

Kuhlthau (1993) prosessimallin perusteella laadittu käyttötarkoituksen mukainen tietotyypijaottelu:

Ideoita tutkimusaiheen valitsemiseen - tietoa joka auttaa löytämään tutkimuksen aiheen tai soveltamiskohteen. Tämän tyypistä tietoa oletetaan haettavan lähinnä tehtävän alussa.

Aiheen valintaa tukevat tulokset - tuloksia, jotka vahvistavat oman jo valitun aiheen mielekkyyden.

Vinkkejä kysymyksenasettelun täsmentämiseen

Rautio ja Suhonen (1981) viittauskäytäntöjen perusteella laadittu käyttötarkoituksen mukainen tietotyypijaottelu :

Teoriatieto - tietoa joka tarjoaa käsitteitä, malleja, määritelmiä, jäsennyksiä ja luokitteluja.

Empiriatieto - tietoa empiirisistä tutkimuksista ja niiden tuloksista.

Menetelmätieto - tietoa tutkimuksissa käytetyistä menetelmistä.

Kuvio 3. Tutkimuksessa käytetyt tietotyypijaottelut.

Tätä tutkimusta varten kehitettiin Kuhlthau (1993) tiedonhankinnan prosessimallin perusteella myös toinen tiedon käyttötarkoituksen mukainen jaottelu, joka perustuu työn etenemisen kannalta eri vaiheissa ilmeneviin konkreettisiin tarpeisiin. Ensinnäkin, tehtävän aloitus- ja valintavaiheissa, käyttäjän oletettiin kaipaavan **ideoita aiheen valitsemiseen**. Hakijalla voi kuitenkin olla jo riittävästi aiempaa tehtävää sivuavaa aihetietämystä, jonka varaan hän voi rakentaa annetun tehtävän aiheen, eikä hän näin ollen tarvitse uusia ideoita. Toiseksi, valinta- ja tunnusteluvaiheessa, käyttäjän oletettiin kaipaavan lähinnä **tuloksia tukemaan aihevalintaa**. Kolmanneksi, erityisesti muotoiluvaiheessa, käyttäjän oletettiin kaipaavan **vinkkejä kysymyksenasettelun täsmentämiseksi**.

Lisäksi tutkimuksessa käytettiin Hakalan ja Vakkarin (Vakkari 2000, 5) tutkimuksessaan sovelta-
maa Raution ja Suhosen (1981, 25) väitöskirjojen lähdeviitteiden tietotyyppien luokitusta mittaamaan hakijoiden tiedon tarvetta käyttötarkoituksen perusteella. Hakalan (1999, 37) mukaan hakijoiden voidaan olettaa tarvitsevan 1) aluksi **teoriatietoa** johdatuksena työn yleiseen aihepiiriin, 2) **empiriatietoa** jäsenyneeemmän aiheheräyksen löytämiseksi tehtävän keskivaiheilla ja 3) **metoditietoa** ongelman ratkaisun suunnittelussa tehtävän lopulla. Tämä tietotyyppien tarvejärjestys ei kuitenkaan

sulje pois sitä, etteikö hakija saattaisi kaivata empiria- tai metoditietoa jo tehtävän alussa tai teoria-tietoa myös tehtävän lopussa.

4.2. Viitteiden hyödyllisyysarvioiden muutos

Tutkimuksessa haluttiin selvittää, miten arviot viitteiden hyödyllisyydestä muuttuvat tehtävän aikana. Vakkarin ja Hakalan (2000, 549) tutkimustulosten perusteella voitiin olettaa, että täysin hyödyllisiksi arvioitujen viitteiden suhteellinen osuus on pienempi tehtävän lopussa kuin alussa ja että kaiken kaikkiaan tehtävän loppupuolella kelpuutetaan vähemmän viitteitä hyödyllisiksi kuin tehtävän alkupuolella.

Käyttäjää pyydettiin arvioimaan viitteitä niiden hyödyllisyyden perusteella, koska heidän arvioitiin ymmärtävän tämän käsitteen helpommin kuin relevanssin käsitteen. Tavoitteena oli, että käyttäjät arvioisivat viitteiden hyödyllisyyttä nimenomaan oman työn kannalta hakuhetkellä. Käyttäjien arvioita hyödyllisyydestä mitattiin asteikolla täysin hyödyllinen, osittain hyödyllinen, mahdollisesti hyödyllinen ja ei lainkaan hyödyllinen. Täysin hyödyllisen kategorian oletettiin vastaavan viitteitä, joiden arvellaan liittyvän osuvasti käyttäjän tarkastelemaan aiheeseen. Osittain hyödyllisiksi oletettiin valittavan viitteitä, jotka liittyvät osittain esillä olevaan ongelmaan tai tarjoavat ongelmaa sivuavaa tietoa. Esimerkiksi osittain hyödylliseksi voidaan valita dokumentti, jossa tarkastellaan relevanttia tutkimusmenetelmää, mutta tutkimuskohde on eri kuin käyttäjän valitsema.

Tätä tutkimusta varten päätettiin käyttää myös mahdollisesti hyödyllisen kategoriaa, koska sen arveltiin erotteluvan osittain hyödyllisiksi katsotuista viitteistä ne, joiden hyödyllisyydestä käyttäjällä ei ole selvää käsitystä. Saracevicin (1996, 214) relevanssijaottelun mukaisesti mahdollinen hyödyllisyys voi myös ilmetä siten, että käyttäjä valitsee viitteen ainoastaan aiheenmukaisuuden perusteella selvittääkseen mahdollisen hyödyn omalle työlleen vasta dokumentin hankittuaan. Esimerkiksi tehtävän alussa käyttäjien voi olettaa arvioivan viitteitä niiden mahdollisen hyödyllisyyden perusteella, kun aiheen tarkastelun kannalta keskeisiä käsitteitä ei vielä tunneta ja kun minkä tahansa aihetta sivuavan viitteen voi arvella auttavan tiedonhankinnassa eteenpäin.

4.3. Viitteiden tietotyyppiarvioiden muutos

Löytyneiden viitteiden osalta haluttiin selvittää, **miten arviot tietotyyppien kuvaavuudesta muuttuvat tehtävän aikana**. Vakkarin (2000, 5) tekemässä tutkimuksessa informaatiotutkimuksen opis-

kelijat arvioivat hyödyllisten viitteiden tarjoavan tehtävän alussa ja keskivaiheilla eniten tausta- ja teorian tietoa, kun taas tehtävän lopulla he arvioivat viitteiden tarjoavan eniten yksityiskohtaista tietoa. Lisäksi tehtävän lopussa haettiin jonkin verran enemmän metodi- ja empiriatietoa kuin tehtävän alussa.

Tässä tutkimuksessa oltiin erityisesti kiinnostuneita vertaamaan tietotyyppi-arvioiden muutosta Vakkarin (2000, 4) tekemiin havaintoihin informaatiotutkimuksen opiskelijoiden tekemistä viitteiden tietotyyppiannista. Myös psykologian opiskelijoilla oletettiin yleisen tiedon tarpeen vähenevän ja metodi-, empiria- ja yksityiskohtaisen tiedon lisääntyvän tehtävän aikana. Viitteiden tietotyyppi-arvioista ei kerätty aineistoa teorian tiedon osalta, mikä huomattiin myöhemmin selväksi puutteeksi. Oletettiin, että ideoita tutkimusaiheen valitsemiseen kaivataan lähinnä tehtävän alussa. Aiheen valintaa tukevia tuloksia ja vinkkejä kysymyksenasettelun täsmentämiseksi oletettiin kaivattavan lähinnä Kuhlthau (1993) prosessimallin muotoiluvaiheessa.

4.4. Haun onnistuminen käyttäjän näkökulmasta

Sun (1994, 216) määritelmiin perustuen tässä tutkimuksessa pyrittiin mittaamaan haun onnistuneisuutta käyttäjän näkökulmasta **kattavuusarvion** ja **kokonaistyytyväisyyden** perusteella. Käyttäjiltä kysyttiin, kuinka kattavasti he uskoivat löytäneensä viitteitä tehtäväänsä varten. Kattavuusarvion oletettiin näin vastaavan käyttäjien käsitystä saannista, siis kuinka tyhjentävästi käyttäjä uskoi löytäneensä hyödyllisiä viitteitä tehtävän sen hetkisen tarpeen mukaisesti. Tyytyväisyydellä pyrittiin mittaamaan käyttäjien tyytyväisyyttä heidän löytämiinsä viitteisiin kokonaisuudessaan. Kokonaistyytyväisyyden oletettiin puolestaan mittaavan käyttäjien käsitystä viitteiden laadusta. Käyttäjien oletettiin olevan tyytyväisiä löydettyihin viitteisiin kokonaisuudessaan, jos he arvioivat viitteiden olevan myös aiheenmukaisia, tarjoavan kaivatun tyyppistä tietoa ja hyödyttävän omaa työtä.

Korrelaation avulla selvitettiin, **onko tyytyväisyyden ja kattavuuden välillä riippuvuutta**. Oletuksena oli, että mitä kattavammin käyttäjät arvioivat löytävänsä viitteitä sitä tyytyväisempiä he ovat saamiinsa viitteisiin. Lisäksi haluttiin selvittää haun onnistuneisuuden mittareiden riippuvuus löytyneiden hyödyllisten viitteiden määrästä, aiheenmukaisuudesta, hyödyllisyydestä sekä tietotyyppiannista. Mikäli käyttäjä arvioi haun kattavuutta viitteiden laadun perusteella, täytyy tyytyväisyyden ja kattavuuden korreloida vaikkei viitteitä löytyisikään paljon. Lisäksi käyttäjän pitäisi olla ainakin jossain suhteessa myös tyytyväinen löytämiinsä viitteisiin, jos hän kokee haun

kattavaksi, sillä käyttäjän ei voi olettaa valitsevan hyödyttömiä viitteitä. Käsitys haun kattavuudesta on siis seurausta sekä laadusta että määrästä. Sen sijaan voidaan olettaa, että käyttäjä on tyytyväinen löytämiinsä viitteisiin kokonaisuudessaan, vaikkei hän löytäisikään niitä paljon ja siis kattavasti.

5. Tutkimuksen kulku ja käytetyt menetelmät

Tutkimukseen osallistui 22 Tampereen yliopiston psykologian opiskelijaa. Opiskelijoiden tehtävänä oli laatia valitsemastaan psykologian alaan liittyvästä aiheesta ensin alustava tutkimussuunnitelma, ja sen jälkeen perusteellisempi lopullinen tutkimussuunnitelma suppeaa empiiristä tutkimusta varten. Tehtävä oli annettu heille psykologian aineopintojen kurssin yhteydessä. Opiskelijoita haastateltiin ja pyydettiin tekemään tehtävänsä liittyviä tiedonhakuja kaksi kertaa, ensin kurssin alkuvaiheessa, vähän ennen ensimmäisen alustavan tutkimussuunnitelmansa luovutusta, ja lopussa, viimeistellyn tutkimussuunnitelman luovutuksen aikoihin. Haastattelut ja tiedonhaut tehtiin Tampereen yliopiston informaatiotutkimuksen laitoksen tiloissa vuoden 2000 syys- marraskuussa. Tutkimuksessa kerättiin aineistoa muun muassa kyselylomakkeiden, hakutallenteiden, strukturoidun teemahaastattelun sekä käyttäjien pitämien hakupäiväkirjojen avulla. Keskeisenä teemana oli kerätä runsaasti sekä **kvantitatiivista** että **kvalitatiivista** aineistoa.

5.1. Aineistonkeruumenetelmät

Tutkimusaineistoa kerättiin tiedonhaku- ja haastattelutilanteissa alku- ja loppuhaastattelujen avulla sekä tallentamalla tiedonhakujen kulku ja hakutulokset. Haastattelut tehtiin pääasiallisesti kyselylomakkeiden avulla, mutta alkuhaastattelun yhteydessä tehtiin myös lyhyt nauhoitettu teemahaastattelu. Haut tallennettiin videonauhalle. Käyttäjien tekemät hakulausekkeet ja heidän löytämät hyödylliset viitteet tulostettiin paperille.

Kyselylomakkeiden avulla pyrittiin mittaamaan kunkin henkilön tiedonhankinnan vaihe, hakukokemus, arviot haetun informaation ja viitteiden tietotyyppien kuvaavuudesta sekä arvio haun onnistuneisuudesta yleensä. Tiedonhankinnan vaiheita mittaava lomake 1 (Liite 1) vastaa pieniä muutoksia lukuun ottamatta Kuhlthaun (1993, 96) Process Survey -lomaketta. Lomakkeen tarkoituksena oli selvittää käyttäjän Kuhlthaun tiedonhankinnan prosessimallin mukainen vaihe. Hakukokemusta mittaava lomake 2 (Liite 2) laadittiin tutkimuksen tarpeita varten, ja sen avulla mitattiin ensimmäisellä hakukerralla mahdollisen OPAC-hakujärjestelmien ja PsychINFO:n käyttökokemuksen vaikutusta tuloksiin. Haetun informaation tietotyyppiä mittaavan lomakkeen 3A (Liite 3)

avulla pyrittiin ensimmäisellä hakukerralla selvittämään, minkä tyyppisen tiedon hakijat odottivat hyödyttävän heitä. Toisella hakukerralla hakujärjestelmän käyttökokemusta kysyttiin ainoastaan PsychINFO:n osalta, ja hakujärjestelmän käyttökokemusta ja haetun informaation tietotyyppisiä mittaavat lomakkeet yhdistettiin samalle lomakkeelle (Liite4). Viitteiden tietotyyppisiä mittaavan lomakkeen 3B (Liite 5) avulla pyrittiin selvittämään, minkä tyyppistä omaa työtä hyödyttävää tietoa dokumenttien arvioitiin viitteiden perusteella antavan omalle työlle. Haun onnistuneisuutta mittaavan lomakkeen 4 (Liite 6) avulla mitattiin haun onnistuneisuutta kokonaisuudessaan kysymällä, miten kattavasti käyttäjät uskoivat löytäneensä viitteitä omaa työtään varten ja kuinka tyytyväisiä he olivat saamiinsa viitteisiin kokonaisuudessaan.

Haetun informaation tietotyypeissä esiintyvä teorian tieto unohdettiin lisätä viitteiden tietotyyppisiä mittaavaan lomakkeeseen 3B, mikä huomattiin jällempäin selväksi puutteeksi. Vakkarin ja Hakalan (2000) tutkimuksen perusteella teorian tiedon osuutta olisi perustellusti voitu kysyä myös viitteiden osalta. Viitteiden tietotyyppisiä mittaavassa lomakkeessa 3B **aiheenmukaisuutta**, **tekijän arvostettavuutta** ja **julkaisun arvostettavuutta** mittaavat kriteerit eivät puolestaan esiinny haetun informaation tietotyypeissä, koska näitä arvioita ei luonnollisesti voi tehdä kuin vasta viitteiden perusteella. Aiheenmukaisuus, tekijän arvostettavuus ja julkaisun arvostettavuus eivät myöskään ole varsinaisia tiedon tyyppisiä, vaikka ne onkin lomakkeen yksinkertaistamiseksi otsikoitu tietotyyppien alle. Lisäksi tekijän ja julkaisun arvostettavuutta sekä muutamia muita eri lomakkeissa mitattuja tekijöitä ei käsitellä tässä työssä lainkaan, koska niiden tarkastelua ei pidetty tarpeellisena tämän työn kannalta.

Hakutallenteet tehtiin nauhoittamalla tietokoneen näyttökuvaa ja hakijoiden ääneen ajattelu videonauhalle. Teemahaastattelussa kysyttiin käyttäjien tehtävää varten tekemien aiempien hakujen tuloksia sekä haastattelun yhteydessä tehtävän haun tavoitteita. Teemahaastattelu tallennettiin ääninauhalle. Lisäksi hakutallenteet ja teemahaastattelut litteroitiin.

5.2. Haastattelujen ja tiedonhakujen kulku

Haastatteluihin oli varattu kaikkiaan aikaa noin kaksi tuntia. Alku- ja loppuhaastatteluun ja lomakkeiden täyttöön oli varattu aikaa noin puoli tuntia kumpaankin. Tiedonhaun annettiin kestää tarvittaessa jonkin verran pidempään mitä hakuun varatun noin tunnin aikana oli mahdollista, mikäli se sopi haastattelijan ja haastateltavan aikatauluun. Ensin henkilöitä pyydettiin täyttämään Prosess Survey -lomake, hakukokemusta mittaava lomake ja haetun informaation tietotyyppisiä

mittaava lomake. Tämän jälkeen tehtiin teemahaastattelu, jonka lopuksi selostettiin seuraavana suoritettavan tiedonhaun osalta, mitä tarkoitetaan ääneen ajattelulla ja miten käyttäjän tulee menetellä ongelmatilanteessa.

Haut tehtiin PsychINFO-tietokannan web-lomakkeiden avulla. Ongelmatilanteissa käyttäjiä opastettiin eteenpäin ja ääneen ajattelua pyrittiin pitämään yllä. Käyttäjien valitsevat viitteet tulostettiin lopuksi arviointia varten ja käyttäjät saivat viitteistä kopiot omaa tehtäväänsä varten.

Haun jälkeen käyttäjiä pyydettiin ensin merkitsemään kunkin viitteen kohdalle arvio viitteen hyödyllisyydestä asteikolla: hyödyllinen, osittain hyödyllinen, mahdollisesti hyödyllinen ja ei lainkaan hyödyllinen. Maksimissaan pyydettiin arvioimaan ensimmäiset 20 viitettä. Sitten kaikki hyödyllisiksi merkityt viitteet (tarkoittaa myös osittain ja mahdollisesti hyödyllisiä) pyydettiin käymään perusteellisemmin läpi ja täyttämään kunkin viitteet osalta viitteiden tietotyyppien arviointilomake.

Lopuksi käyttäjät saivat vielä täytettäväkseen haun evaluointilomakkeen ja yhden viitteiden tietotyyppien mittaavan lomakkeen jokaista löytynyttä hyödyllistä viitettä kohti. Mikäli viitteiden tietotyyppien arviointilomakkeiden täyttämiseen varattu aika oli selvästi riittämätön annettiin ne kotona täytettäväksi ja palautuksesta annettiin erilliset ohjeet.

Ensimmäinen haastattelukerta pyrittiin tekemään heti tutkimussuunnitelman laadintaa vaativan kurssin alkuun ja toinen haastattelu kurssin loppuun vähän ennen tutkimussuunnitelmien luovutusta ja esitystä. Ensimmäinen haastattelu onnistuttiin tekemään jotakuinkin aikataulun mukaisesti, mutta toisella haastattelukerralla kuusi opiskelijaa oli ehtinyt jo luovuttaa työnsä tai työn luovutusajankohta oli niin lähellä, että hakijat eivät välttämättä olleet enää kovin motivoituneita tehtävänsä suhteen. Opiskelijoiden tehtävä kuitenkin jatkui vielä tutkimussuunnitelman luovutuksen jälkeenkin, sillä seuraavalla kurssilla heidän piti toteuttaa suunnittelemansa tutkimus. Näin ollen osa suunnitelmansa jo luovuttaneista vaikutti edelleen motivoituneilta tekemään ainakin tarkistavia tiedonhakuja.

5.3. *Datan käsittely tarkastelua varten*

Kyselylomakkeiden avulla kerätty data koodattiin tilastollista tarkastelua varten. Lisäksi koodattiin hakutallenteiden avulla saadut määrälliset muuttujat, kuten hakujen aikana löytyneiden viitteiden määrät hyödyllisyysasteittain. Tutkimusaineistoa kerättiin myös muun muassa hauissa käytettyjen

operaattoreiden ja taktiikoiden käytöstä. Tässä työssä kuitenkin tarkastellaan vain **tietotyypiarvioiden, hyödyllisyysarvioiden** sekä **haun evaluoinnin** muutosta tehtävän aikana. Kyselylomakkeet oli laadittu suurimmalta osin siten, että vastausvaihtoehdot oli valmiiksi luokiteltu järjestysasteikollisiksi.

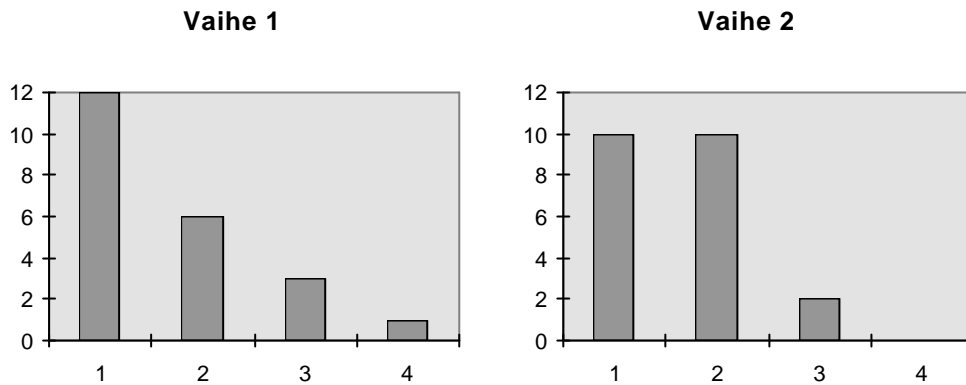
Kuhlthaun (1993) tiedonhankinnan prosessimallin mukaista vaihetta mittaava Process Survey -lomake oli sen sijaan laadittu monivalintakysymysten muotoon, joten tutkijoiden oli ensin analysoitava tulokset. Kuhlthaun mallissa oletetaan erilaisten tuntemusten, ajatusten ja tekojen ilmenemisen ensisijaisesti jossain tietyssä tehtävän suorituksen vaiheessa. Process Survey -lomakkeen vastausvaihtoehdot on laadittu perustuen tähän oletukseen, minkä vuoksi kunkin vastausvaihtoehdon oletetaan viittaavan tiettyyn prosessimallin vaiheeseen. Vastausvaihtoehdot voivat kuitenkin kuvata useampaakin tehtävän vaihetta, joten tulokset oli ensin analysoitava kvalitatiivisesti, jotta voitaisiin määrittää, missä vaiheessa kukin hakija todennäköisimmin oli. Tässä tutkimuksessa kolme tutkijaa kävi ensin itsenäisesti läpi kaikki Process Survey -lomakkeet jakaen vastaajat oman arvionsa perusteella Kuhlthaun luokituksen mukaisiin vaiheisiin. Tämän jälkeen tutkijat vertasivat tuloksia. Mikäli tulkinnat olivat eriäviä, niistä keskusteltiin ja haettiin yhteinen ratkaisu. Vertailun tueksi Carol Kuhlthaulta saatiin myös hänen käyttämänsä vastausvaihtoehtojen vaiheluokitus.

5.4. Tilastollisten analyysimenetelmien valinta

Tulosten analysoinnissa käytetyt tilastolliset analyysimenetelmät pyrittiin valitsemaan tarkoitukseen sopiviksi ottaen huomioon tutkimuksen suhteellisen pienen otoskoon. Tarkastelen työssäni myös pääasiassa järjestysasteikollisia muuttujia, minkä vuoksi keskilukujen esittämisessä olisi suositeltavaa käyttää mediaaneja tai moodeja (ks. mm. Yli-Luoma 1997, 40). Olen kuitenkin päätenyt esittämään keskiluvut keskiarvoina, koska keskiarvolla voidaan tarkemmin ilmaista keskiluvun sijoitumista järjestysasteikon luokkien välillä, etenkin kun luokkia on vähän. Esimerkiksi kuviossa 4 on esitetty erään tutkimuksessa mitatun muuttujan⁹ jakaumat kahdessa vaiheessa mitattuna. Ensimmäisellä kerralla muuttujan mediaaniksi saatiin 1 ja keskiarvoksi 1,7. Toisessa vaiheessa mediaaniksi saatiin 2 ja keskiarvoksi 1,6. Mediaanin perusteella ero vaiheiden välillä näyttää siis paljon suuremmalta kuin keskiarvolla mitattuna. Kuvion 4 jakaumia tarkasteltaessa voidaan kuitenkin tulla siihen tulokseen, että luokan 1 äänimäärän muutos vaiheiden välillä on vain 2 ääntä, eli juuri

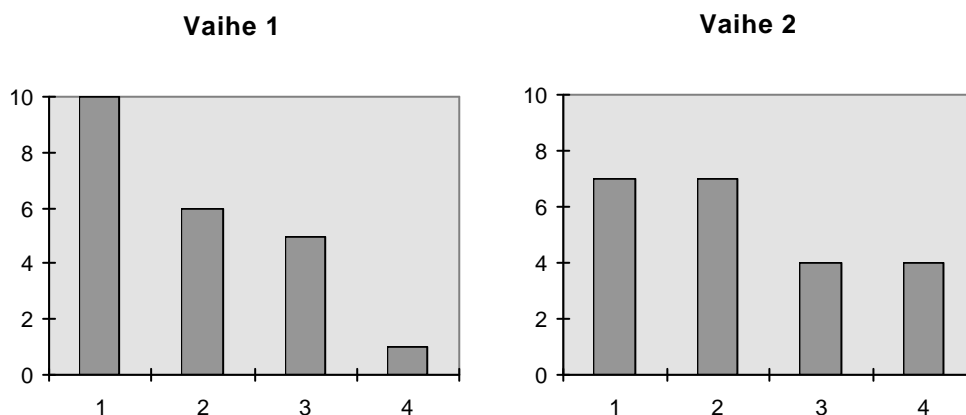
⁹ Teoriatiedon odotettu anti ennen tiedonhakuja, mitattuna tehtävän eri vaiheissa.

sen verran kuin mediaanin vaihtumiseksi tarvitaan. Toisaalta luokan 2 äänimäärä on kasvanut kaikkiaan neljällä äänellä, joten muuttuja on toisessa vaiheessa jopa aavistuksen enemmän painottunut luokille 1 ja 2 yhteensä kuin ensimmäisessä vaiheessa. Siten keskiarvo on molemmilla kerroilla lähes sama.



Kuvio 4. Teoriatiedon odotettu anti tehtävän alussa (Vaihe 1) ja lopussa (Vaihe 2).

Vastaavasti kuviossa 5 erään toisen mitatun muuttujan¹⁰ mediaaniksi saatiin molemmilla kerroilla 2, kun taas keskiarvo oli ensimmäisessä vaiheessa 1,9 ja toisessa 2,2. Tällöin mediaani ei tee eroa vaiheiden välillä, mutta keskiarvon perusteella jakauman voimakkaampi vinous ensimmäisessä vaiheessa tulee paremmin esille. Vähäluokkaisten asteikkojen suhteen mediaani siis joko liioittelee tai vähättelee eroa, kun taas keskiarvo näyttää ilmaisevan tarkemmin niin jakauman vinouden kuin myös jakauman keskiluvun sijoittumisen luokkien välille.



Kuvio 5. Aiheen täsmennystä auttavan tiedon odotettu anti mitattuna tehtävän alussa (Vaihe 1) ja lopussa (Vaihe 2).

¹⁰ Aiheen täsmennystä auttavan tiedon odotettu anti ennen tiedonhakuja, mitattuna tehtävän eri vaiheissa.

Populaatiolle ei voida esittää yhtä selkeätä mielipidettä, jos keskiarvo osuu järjestysasteikollisen muuttujan luokkien välille, kuten moodin tai mediaanin avulla on mahdollista. Kuvioiden 4 ja 5 esimerkkien kaltaisten neliluokkaisten muuttujien kohdalla keskiarvoa pitää mielestäni lukea siten, että keskiarvo useimmiten ilmaisee asteikon sen kohdan, jonka viereiset luokat ovat saaneet yhteensä eniten kannatusta. Keskiarvo ei kuitenkaan ilmaise muuttujan vinouden¹¹ suuntaa tai määrää. Keskiarvon ohella voi käyttää myös vinouslukua¹² ilmaisemaan vähintään muuttujan asteikon vinouden suunnan. Tutkimuksessani olen kuitenkin katsonut paremmaksi verrata ainoastaan keskilukuja keskenään kuin laatia yksityiskohtaisia esityksiä jakaumien vinoudesta. Mikäli jakauma on sitä vastoin osoittautunut erittäin epätasaiseksi, olen esittänyt siitä tarkemman analyysin.

Eri vaiheissa mitattujen jakaumien välisen eron merkitsevyyttä on puolestaan testattu tilastollisesti kahdella eri testillä riippuen siitä, onko havaintoyksikkönä hakijat vai heidän löytämänsä viitteet. Kun havaintoyksikkönä olivat hakijat, verrattiin saman muuttujan suhteen kahdessa eri vaiheessa mitattuja arvoja Wilcoxonin kahden rinnasteisen¹³ otoksen testillä. Viitteiden vertaamiseksi käytettiin Mann-Whitneyn kahden riippumattoman¹⁴ otoksen *U*-testiä. Molemmat testit ovat ei-parametrisiä, mikä onkin muun muassa Cramerin (1994, 53) perusteella paikallaan, kun muuttujat on luokiteltu järjestysasteikollisesti ja jakaumat ovat vinoja. Niin Wilcoxonin kuin Mann-Whitneyn *U*-testin osalta on myös hyvä huomata, etteivät ne perustu keskiarvojen vertailuun. Wilcoxonin testissä verrataan, miten merkitsevästi otoksen kunkin havaintoyksikön saama muuttujan arvo poikkeaa eri aikana mitatusta saman havaintoyksikön muuttujan arvosta (ks. Cramer 1994, 114). *U*-testissä taas mitataan, miten monta kertaa korkeammalle tai alemmalle sijalle otoksen muuttujan arvot sijoittuvat verrattuna toiseen otokseen (ks. Cramer 1994, 104).

Korrelaatioiden laskemiseen käytettiin 'Kendallin tau' -järjestyskorrelaatiota. Viitteiden tietotyyp-
piarvioille tehdyssä faktoroinnissa (ks. Luku 6.2.; Taulukko 3) käytettiin pienimmän neliösumman faktorointia. Viitteiden saamien eri ominaisuuksien ja haun onnistuneisuuden kriteerien välisten riippuvuuksien laskemisessa käytettiin kuvion 6 mukaista osittaiskorrelaation kaavaa.

¹¹ Jakauma on vino, jos keskiarvo ja mediaani sijoittuvat eri kohtiin jakaumaa.

¹² Vinousluku = moodin ja keskiarvon erotus jaettuna keskihajonnalla.

¹³ Populaatio on molemmissa otoksissa sama ja mittauksessa on käytetty joko samaa muuttujaa (esim. eri ajankohtana mitattuna) tai kahta samoin luokiteltua muuttujaa.

¹⁴ Populaatio voi olla molemmissa otoksissa eri, mutta mittauksessa on käytetty joko samaa muuttujaa (esim. eri ajankohtana mitattuna) tai kahta samoin luokiteltua muuttujaa.

$$r_{xy,z} = \frac{r_{xy} - (r_{xz})(r_{yz})}{\text{sqrt}[(1-r_{xz}^2)(1-r_{yz}^2)]}$$

Kuvio 6. Tutkimuksessa käytetty kaava osittaiskorrelaation laskemiseksi.

Lowryn (1999-2000, verkkojulkaisu) mukaan osittaiskorrelaation avulla voidaan laskea kahden muuttujan välisen riippuvuuden suuruus, kun kolmannen, molempien muuttujien kanssa korreloivan muuttujan, vaikutus suljetaan pois eli vakioidaan. Kun osittaiskorrelaatiossa käytetään kuvion 6 kaavaa, on Z:n suhteen vakioitu X:n ja Y:n välinen osittaiskorrelaatio yleensä pienempi kuin X:n ja Y:n vakioimaton korrelaatio. Mikäli ero on vain pieni, voidaan sanoa, ettei Z:lla ole juurikaan vaikutusta X:n ja Y:n väliseen riippuvuuteen. Jos X:n ja Y:n riippuvuus puolestaan heikkenee huomattavasti osittaiskorreloinnin tuloksena, on Z:n vaikutus ilmeinen.

6. Tulokset

Tässä työssä keskitytään tarkastelemaan, minkä tyyppistä tietoa käyttäjät odottivat tarvitsevansa ja minkä tyyppistä tietoa he arvioivat viitteiden perusteella löytäneensä tehtävän eri vaiheissa. Lisäksi tarkastellaan viitteiden hyödyllisyysarvioiden, aiheenmukaisuuden ja haun onnistuneisuusarvioiden muutosta tehtävän eri vaiheissa. Tutkimuksen yhtenä tavoitteena oli myös pyrkiä valitsemaan hakutaidoiltaan ja hakujärjestelmäkokemukseltaan mahdollisimman homogeeninen ryhmä käyttäjiä. Kaikkien käyttäjien todettiin olevan hakukokemukseltaan ja taidoiltaan jokseenkin samalla tasolla eikä hakutaitojen ja järjestelmätuntemuksen myöskään havaittu vaikuttaneen merkittävästi käyttäjien odotuksiin tai hakujen tuloksiin. Tässä työssä ei sen vuoksi käsitellä yksityiskohtaisemmin OPAC ja PsychINFO -viitetietokantojen käyttökokemuksen ja hakutaitojen yhteyttä psykologian opiskelijoiden osalta.

6.1. Tietotyyppien odotettu tarve

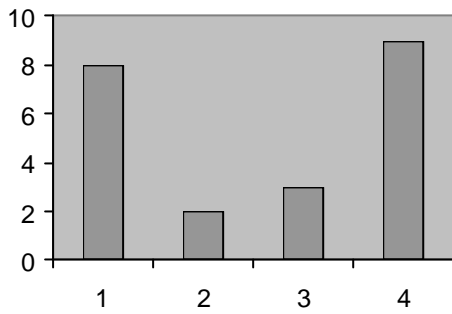
Ennen hakuja opiskelijoita pyydettiin arvioimaan missä määrin he odottivat eri tyyppisen tiedon vastaavan heidän sen hetkistä tiedon tarvettaan. Taulukossa 1 on esitetty tietotyyppien saamat keskiarvot tehtävän alussa (Vaihe I) ja sen lopussa (Vaihe II).

Taulukko 1. Tietotyyppien odotettu tarve ennen tiedonhakuja: keskiarvot hakukerroittain per henkilö (tietotyyppi kuvaa 1= osuvasti, 2= jossain määrin, 3= heikosti, 4= ei lainkaan) sekä eri vaiheitten välisen eron tilastollinen merkitsevyys.

Tietotyypit	Vaihe I	Vaihe II	p
Teoriatietoa	1,7	1,6	.825
Yleistä taustatietoa	1,3	2,3	.001
Yksityiskohtaista tietoa	2,8	1,8	.004
Menetelmätietoa	2,1	1,6	.080
Empiriatietoa	2,1	1,5	.013
Ideoita aiheen valitsemiseen	2,6	3,8	.001
Vinkkejä ongelman täsmentämiseen	1,9	2,2	.096
Tuloksia tukemaan aiheenvalintaa	2,2	2,4	.308
<i>n</i>	22	22	

Ensimmäisellä hakukerralla opiskelijat arvioivat yleisen taustatiedon kuvaavan heidän tiedontarvettaan kaikkein parhaiten ($ka=1,3$). Haun avulla mahdollisesti löytyvän kirjallisuuden odotettiin tarjoavan myös teoriatietoa ($ka=1,7$) sekä antavan vinkkejä kysymyksenasettelun täsmentämiseksi ($ka=1,9$). Kaikkein vähiten odotettiin yksityiskohtaista tietoa ($ka=2,8$) ja ideoita aiheen valitsemiseen ($ka=2,6$). Toisella hakukerralla kaikkein parhaiten sen hetkistä tiedontarvetta kuvaaviksi tietotyypeiksi arvioitiin empiriatieto ($ka=1,5$), menetelmätieto ($ka=1,6$), teoriatieto ($ka=1,6$) ja yksityiskohtainen tieto ($ka=1,8$). Ideoita aiheen valitsemiseen kaivattiin odotetusti enää hyvin vähän ($ka=3,8$).

Yleistä taustatietoa ja ideoita aiheen valitsemiseen otaksuttiin tarvittavan merkitsevästi vähemmän tehtävän lopussa kuin alussa ($p=.001$). Samoin vinkkejä ongelman täsmentämiseen arveltiin tarvittavan lopussa hieman vähemmän ($p=.096$). Sen sijaan yksityiskohtaisen tiedon otaksuttu tarve lisääntyi merkitsevästi ($p=.004$) ja empiriatiedon tarve melkein merkitsevästi ($p=.013$) ensimmäisestä toiseen hakuun. Myös menetelmätiedon tarve lisääntyi jossain määrin tehtävän aikana ($p=.080$). Teoriatiedon odotettiin olevan tärkeää niin tehtävän alussa kuin lopussakin, mutta aiheen valintaa tulevia tuloksia kaivattiin tehtävän eri vaiheissa melko vähän. Tulokset ovat yhtenevät Kuhlthaun (1993) tiedonhankinnan prosessimallin ja Vakkarin (2000, 5) tutkimustulosten kanssa.



Kuvio 7. Aiheenvalintaa auttavien ideoiden odotettu anti henkilöittäin ensimmäisellä haastattelukerralla ($ka=2,6$; $N=22$).

Vaikka tehtävän alussa hakijat olivat keskiarvon perusteella jo valinneet aiheensa, painottuivat vastaukset kuitenkin selvästi ääriarvoihin: 9 opiskelijaa (41%) arvioi, ettei lainkaan odota löytävänsä ideoita aiheen valitseen, kun taas 8 hakijaa (36%) arvioi tietotyypin kuvaavan heidän tiedontarvettaan osuvasti (Kuvio 7). Kuhlthauin process survey -lomakkeen perusteella ideoita aiheen valintaan kaipasivat yhtä lailla niin aloitus-, valinta- kuin tunnusteluvaiheessakin olevat. Tehtävän alussa aihetietämyksessä pitemmälläkin olevat näyttivät siis olevan avoimia uusille aiheenvalinnoille. On huomattava, että tehtävän suorittaminen saattoi olla joillekin kiinteämmin sidoksissa aiempiin opintoihin tai kiinnostuksen aiheisiin, minkä vuoksi heillä saattoi olla myös aiempaa tietämystä menetelmistä tai teorioista vaikkei varsinaista soveltamiskohdetta ollut vielä valittu. Ideoita aiheen valintaan eivät kuitenkaan kaivanneet ne, jotka ilmaisivat hakevansa yksityiskohtaista tietoa. Negatiivinen korrelaatio tietotyyppien **ideoita aiheen valitsemiseen** ja **yksityiskohtaista tietoa** välillä oli merkitsevä ($r = -.53$; $p = .011$)

Myös menetelmätiedon osalta jakauma oli tehtävän alussa lievästi ääripäihin painottuva: 32% arvioi menetelmätiedon kuvaavan heidän tiedontarvettaan osuvasti, 27% jossain määrin ja 41% heikosti (ei lainkaan = 0%). Menetelmätiedon tarve on yllättävänkin suuri, kun ottaa huomioon informaatiotutkimuksen opiskelijoilla tehdyn vastaavan tutkimuksen (Hakala 1999, liite 7). Hakala ei tutkimuksessaan varsinaisesti tarkastele ennen hakuistuntoja esitettyjen odotusten osuutta, mutta aineiston perusteella ainoakaan 11:stä informaatiotutkimuksen opiskelijasta ei tuolloin ilmaissut kaipaavansa menetelmätietoa aivan tehtävänsä alussa. Tässä tarkasteltavilla psykologian opiskelijoilla oli sen sijaan tehtävänään laatia tutkimussuunnitelma nimenomaan kokeellista tutkimusta varten, joten tämä selittää ainakin osittain yksityiskohtaiseksi luonnehdittavan menetelmätiedon tarpeen jo heti tehtävän alussa.

6.2. Tietotyypeittäin arvioitu viitteiden anti

Viitetietokantojen hakujärjestelmät tarjoavat erilaisia mahdollisuuksia hakutulosten tarkasteluun. Myös viitteiden sisältämät dokumenttien kuvaustavat vaihtelevat. PsycInfo -tietokannassa viitteiden bibliografiset kuvaukset sisältävät muun muassa lyhyen tiivistelmän dokumenttien aiheesta ja sisällöstä. Hakujärjestelmä mahdollistaa viitteiden yksittäisen valitsemisen haun jälkeen tapahtuvaa tarkastelua varten (esim. tallennettavaksi, sähköpostiliitteenä lähetettäväksi tai paperitulostusta varten). Tutkimukseen osallistuneita opiskelijoita neuvottiin merkitsemään haun aikana kaikki heitä kiinnostavat viitteet tulostettavaksi hakuistunnon jälkeen tapahtuvaa yksityiskohtaisempaa relevanssi- ja tietotyyppi-arviointia varten. Tulostettujen viitteiden hyödyllisyyttä pyydettiin ensin arvioimaan neliportaisella asteikolla: **hyödyllinen, osittain hyödyllinen, mahdollisesti hyödyllinen, ei lainkaan hyödyllinen**. Löytyneiden relevanttien viitteiden hyödyllisyyttä omalle työlle pyydettiin edelleen arvioimaan käyttämällä samaa tietotyyppiluokitusta kuin kysyttäessä hakuodotuksia. Epärelevanttien viitteiden osalta tietotyyppi-arvioita ei tehty. Ennen viitteiden tietotyyppi-arvioita tehtyjä hyödyllisyysarvioita tarkastellaan luvussa 6.4.

Viitteitten arvioinnissa ei kysytty **teoriatiedon** mahdollista antia, mikä huomattiin jälkikäteen ilmeiseksi puutteeksi. Tietotyyppien kuvaavuuden lisäksi hakijoita pyydettiin arvioimaan viitteiden aiheenmukaisuutta. Kyselylomakkeessa kehoitettiin tarkastelemaan tietotyyppi-arvioita ja aiheenmukaisuutta tehtävän sen hetkisen vaiheen kannalta.

Taulukko 2. Löytyneiden hyödyllisten viitteiden arvioidun tietotyyppiannin ja aiheenmukaisuuden keskiarvot laskettuna viitteille (1= kuvaa osuvasti, 2= jossain määrin, 3= heikosti, 4= ei lainkaan).

Tietotyyppit	Vaihe I	Vaihe II	p
Yleistä taustatietoa	1,9	2,2	.007
Yksityiskohtaista tietoa	1,8	2,1	.038
Menetelmätietoa	2,3	2,5	.029
Empiirisiä tutkimustuloksia	1,9	2,1	.356
Ideota aiheen valitsemiseen	2,4	3,2	.000
Vinkkejä ongelman muotoilemiseen	2,3	2,8	.000
Tuloksia tukemaan aiheenvalintaa	2,2	2,6	.004
Viitteiden arvioitu aiheenmukaisuus	1,6	1,8	.151
<i>n</i>	150	98	

Arviot tietotyyppien kuvaavuudesta olivat ensimmäisellä hakukerralla keskimäärin positiivisempia kuin toisella kerralla. Käyttäjät arvioivat viitteiden sisältävän tehtävän alussa osuvammin kaikkia mitattuja tietotyyppiä kuin tehtävän lopussa. Ero vaiheiden välillä oli merkitsevä yleisen

taustatiedon ($p=.007$), uusien aiheenvalintaa tukevien ideoiden ($p=.000$), ongelman muotoilussa auttavien vinkkien ($p=.000$) ja aiheenvalintaa tukevien tulosten ($p=.004$) osalta. Yksityiskohtaisen tiedon ($p=.038$) ja menetelmätiedon ($p=.029$) osalta ero oli melkein merkitsevä. Molemmilla hakukerroilla viitteiden arvioitiin tarjoavan parhaiten yksityiskohtaista tietoa, yleistä taustatietoa ja empiiristä tietoa. Molemmilla kerroilla arvioitiin viitteiden tarjoavan vähiten uusia ideoita aiheen valitsemiseen. Viitteet arvioitiin ensimmäisellä hakukerralla hieman aiheenmukaisemmiksi ($k_a=1,6$) kuin toisella hakukerralla ($k_a=1,8$).

Taulukko 3. Viitteiden tietotyypin faktorit eri hakukerroilla (faktorien ominaisarvo > 1).

Tietotyypit	Vaihe I (n=138)		Vaihe II (n=96)	
	Faktori 1	Faktori 2	Faktori 1	Faktori 2
Yleistä taustatietoa	,554	,023	,462	-,068
Yksityiskohtaista tietoa	,404	,600	,211	,625
Menetelmätietoa	,145	,482	-,052	,408
Empiirisiä tutkimustuloksia	,046	1,014	,028	,810
Ideota aiheen valitsemiseen	,800	,193	,570	,053
Vinkkejä ongelman muotoilemiseen	,799	,263	,833	,208
Tuloksia tukemaan aiheenvalintaa	,622	,381	,521	,518

Koska viitteiden arvioissa esiintyi useita merkitseviä tai melkein merkitseviä korrelaatioita tietotyyppien välillä (Liite 7), tietotyyppien välisten suhteiden selittämiseksi laskettiin myös faktorit (Taulukko 3). Faktorianalyysissä käytettiin pienimmän neliösumman faktorointia ja tarkasteluun valittiin faktorit, joiden ominaisarvo on suurempi kuin yksi. Ensimmäisellä kerralla viitteet antoivat aiheenvalinnassa ja kysymyksenasettelussa epävarmoille ideoita sekä yleistä taustatietoa että yksityiskohtaista tietoa (Vaihe I, faktori 1). Aiheenvalinnassa ja kysymyksenasettelussaan varmemmat valitsivat viitteitä, jotka antoivat lähinnä yksityiskohtaista, empiiristä ja menetelmätietoa (Vaihe I, faktori 2). Toisella hakukerralla joidenkin opiskelijoiden aiheenvalinnan ja kysymyksenasettelun epävarmuus heijastui edelleen yleisen taustatiedon, aiheen valintaa tukevien ideoiden ja ongelman muotoilussa auttavien vinkkien valintana (Vaihe II, faktori 1). Sen sijaan ne, jotka olivat jo tehneet tarkan aiheenrajauksen, valitsivat viitteitä, joiden he arvioivat antavat empiriatietoa, yksityiskohtaista tietoa ja menetelmätietoa (Vaihe II, faktori 2). Tulokset tukevat ja täsmentävät Kuhlthaun (1993) tiedonhankinnan prosessimallia, jonka mukaan tiedon tarve muuttuu yleisestä yksityiskohtaiseksi.

6.3. Tietotyyppiodotusten ja viitteiden tietotyyppiarvioiden vastaavuus

Jotta löydettyjen viitteiden tietotyyppi-arvioita voitaisiin verrata odotusten keskiarvoihin, on taulukossa 4 laskettu ensin viitteiden tietotyyppi-arvioiden keskiarvot henkilöittäin ja sitten näin saaduista arvoista on laskettu keskiarvo koko populaatiolle.

Taulukko 4. Tietotyyppiodotusten (odot.) sekä viitteiden arvioidun annin keskiarvot per henkilö (1= kuvaa osuvasti, 2= jossain määrin, 3= heikosti, 4= ei lainkaan) sekä niiden välisen eron tilastollinen merkitsevyys vaiheittain.

Tietotyypit	Vaihe I			Vaihe II		
	Odot.	Anti	p	Odot.	Anti	p
Yleistä taustatietoa	1,3	2,1	.001	2,3	2,4	.794
Yksityiskohtaista tietoa	2,8	1,9	.013	1,8	1,9	.640
Menetelmätietoa	2,1	2,4	.153	1,6	2,3	.025
Empiriatietoa	2,1	2	1.000	1,5	2	.052
Ideoita aiheen valitsemiseen	2,6	2,4	.404	3,8	3,4	.018
Vinkkejä ongelman täsmentämiseen	1,9	2,3	.105	2,2	3,1	.004
Tuloksia tukemaan aiheenvalintaa	2,2	2,4	.394	2,4	2,8	.223
<i>n</i>	22	21		22	19	

Ensimmäisellä hakukerralla viitteiden arvioitiin tarjoavan merkitsevästi vähemmän yleistä taustatietoa ($p=.001$), kuin opiskelijat ennen hakuja odottivat tarvitsevansa. Menetelmätietoa ($p=.153$) ja vinkkejä ongelman täsmentämiseen ($p=.105$) löydettiin jonkin verran odotettua vähemmän. Sen sijaan yksityiskohtaista tietoa löydettiin merkitsevästi enemmän ($p=.013$). Vaikka käyttäjät eivät siis odottaneet löytävänsä yksityiskohtaista tietoa tehtävän alussa, he kuitenkin tunnistivat yksityiskohtaiseksi tulkitun tiedon hyödylliseksi sellaista löytäessään.

Aiheen valintaa tukevien ideoiden osalta on huomattava odotusten voimakas jakautuminen tehtävän alussa. Tarkasteltaessa ainoastaan niitä hakijoita, jotka arvioivat tarvitsevansa ideoita aiheen valitsemiseen osuvasti tai jossain määrin, saadaan heidän odotustensa keskiarvoksi 1,2 ja viitteiden annin keskiarvoksi 2,1. Tällöin myös ero vaiheiden välillä muodostuu melkein merkitseväksi ($p=.026$). Hakijat, jotka puolestaan otaksuivat tarvitsevansa uusia ideoita vain heikosti tai ei lainkaan ($ka=3,8$), arvioivat viitteiden perusteella löytäneensä ideoita odotettua enemmän ($ka=2,7$). Tilastollisesti ero on merkitsevä ($p=.009$). Muiden tietotyyppien osalta viitteiden anti vastasi odotuksia.

Tehtävän lopussa viitteiden arvioitu anti vastasi yleisen taustatiedon, yksityiskohtaisen tiedon ja aiheen valintaa tukevien tulosten osalta opiskelijoiden ennen hakuja ilmaisemaa tiedon odotettua

tarvetta. Menetelmätietoa ($p=.025$) ja empiriatietoa ($p=.052$) löytyi puolestaan selvästi odotettua vähemmän. Vinkkejä ongelman täsmentämiseen ($p=.004$) löydettiin merkitsevästi vähemmän. Mahdollisesti käyttäjät epäonnistuivat hakujen muotoilussa, minkä vuoksi ideoita aiheen valitsemiseen löytyi toisella hakukerralla odotettua enemmän ($p=.018$), vaikkei uusia ideoita keskiarvojen perusteella varsinaisesti tarvittu.

6.4. Löytyneiden viitteiden hyödyllisyys

Tiedonhauissa tulostettaviksi valittujen viitteiden hyödyllisyyttä pyydettiin arvioimaan loppuhaastatteluissa asteikolla: 1) täysin hyödyllinen, 2) osittain hyödyllinen, 3) mahdollisesti hyödyllinen ja 4) ei lainkaan hyödyllinen. Edelleen yksityiskohtaisemmin tarkasteltavaksi valittiin ainoastaan täysin hyödyllisiksi, osittain hyödyllisiksi ja mahdollisesti hyödyllisiksi arvioidut viitteet, sillä tutkimuksessa oli ensisijaisesti kiinnostuttu hyödyllisyyden asteista – ei siis epärelevanttien mahdollisista asteista tai epärelevanttien viitteiden määrästä tai osuudesta¹⁵. Taulukossa 5 on esitetty viitteiden määrien keskiarvot sekä prosentuaaliset osuudet hyödyllisyysasteittain henkilö molemmilla hakukerroilla.

Taulukko 5. Viitteiden määrien keskiarvot ja prosentuaalinen osuus hyödyllisyysasteittain eri hakukerroilla sekä erojen tilastollinen merkitsevyys vaiheitten välillä.

Hyödyllisyyden aste	Vaihe I (n=150)		Vaihe II (n=98)		p
	Ka.	%	Ka.	%	
Täysin hyödyllinen	1,8	26	0,9	21	.053
Osittain hyödyllinen	1,9	28	1,5	32	.523
Mahdollisesti hyödyllinen	3,1	46	2,1	47	.149
Hyödylliset yhteensä	6,8	100	4,4	100	.097

Ensimmäisellä hakukerralla opiskelijat löysivät yhteensä 150 ja toisella kerralla 98 relevanttia viitettä. Tehtävän alussa opiskelijat löysivät keskimäärin puolet enemmän relevantteja viitteitä ($ka=6,8$) kuin lopussa ($ka=4,4$). Ensimmäisellä hakukerralla lähes puolet (n. 46%) viitteistä arvioitiin mahdollisesti hyödyllisiksi ($ka=3,1$). Täysin hyödyllisiksi ($ka=1,8$) ja osittain hyödyllisiksi

¹⁵ Loppukäyttäjien epärelevantteiksi arvioimien viitteiden laskeminen on myös vaikeaa. Hakijat saattavat vaihtaa tulostettavaksi tarkoitettujen viitteiden merkintöjä kesken hakujen eivätkä he useinkaan selaa läpi kaikkia hakusettejä. Hakijan kannalta epärelevantteja ovat siten myös kaikki ne viitteet, jotka tiedonhaun aikana selattaessa käydään läpi, mutta joita ei valita myöhemmin tarkasteltavaksi. Epärelevanttien viitteiden kokonaismäärän voisi tosin laskea hakutallenteiden avulla (videonauha ja ääneen ajattelu).

($k_a=1,9$) hyväksyttiin keskimäärin yhtä monta viitettä. Tehtävän lopussa täysin hyödyllisiksi hyväksyttiin huomattavasti vähemmän viitteitä ($k_a=0,9$) kuin tehtävän alussa, ja ero hakukertojen välillä oli tilastollisesti melkein merkitsevä ($p=.053$). Osittain hyödyllisiksi arvioitiin toisella kerralla keskimäärin vain hieman vähemmän viitteitä ($k_a=1,5$; $p=.523$) kuin ensimmäisellä kerralla. Mahdollisesti hyödyllisiä viitteitä löytyi toisella kerralla selvästi vähemmän ($k_a=2,1$; $p=.149$), mutta prosentuaalinen osuus kaikista relevanteista viitteistä pysyi lähes samana ($n. 47\%$).

Tulokset ovat osittain samansuuntaiset kuin vastaavassa informaatiotutkimuksen opiskelijoilla tehdyssä tutkimuksessa (Hakala 1999, 51; Vakkari ja Hakala 2000, 549). Kaikkien hyödyllisten viitteiden määrä väheni miltei samassa suhteessa molemmilla opiskelijaryhmillä. 11 informaatiotutkimuksen opiskelijaa löysi tehtävän alussa yhteensä 86 relevanttia viitettä. Viitteitä löydettiin selvästi vähemmän tehtävän keskivaiheilla (yht. 53 kpl:tta) ja tehtävän lopussa (yht. 63 kpl:tta). Täysin hyödyllisten viitteiden määrä oli jälkimmäisillä hakukerroilla alle puolet siitä mitä tehtävän alussa löydettiin. Kaikista relevanteista viitteistä arvioitiin täysin hyödyllisiksi 59% tehtävän alussa, 41% tehtävän keskivaiheilla ja 35% tehtävän lopussa. On huomattava, että informaatiotutkimuksen opiskelijoilta ei kysytty viitteiden mahdollista hyödyllisyyttä, joten prosentuaaliset odotukset eri populaatioiden välillä eivät ole suoraan verrannollisia. Hyödylliseksi katsottujen viitteiden kokonaismäärän väheneminen ja täysin hyödyllisten viitteiden osuuden puolittuminen tehtävän aikana vastaa kuitenkin tässä tutkimuksessa saatuja havaintoja.

Taulukko 6. Korrelaatiot viitteiden arvioidun hyödyllisyyden, tietotyyppiannin ja aiheenmukaisuuden välillä (* $p<.010$; * $p<.050$).

Viitteiden arvioitu tietotyyppi	Viitteiden hyödyllisyys	
	Vaihe I	Vaihe II
Yleistä taustatietoa	,32**	,27**
Yksityiskohtaista tietoa	,27**	,30**
Menetelmätietoa	,17*	,13
Empiirisiä tutkimustuloksia	,16*	,14
Ideoita aiheen valitsemiseen	,18*	,22*
Vinkkejä ongelman muotoilemiseen	,25**	,42**
Tuloksia tukemaan aiheenvalintaa	,35**	,39**
Viitteiden arvioitu aiheenmukaisuus	,52**	,30**
<i>n</i>	150	98

Arvio hyödyllisyyden asteesta korreloi molemmilla kerroilla merkitsevästi niin aiheenmukaisuuden kuin useimpien tietotyyppiarvioidenkin kanssa (Taulukko 6). Korrelaatio hyödyllisyyden ja aiheenmukaisuuden välillä oli molemmilla kerroilla tilastollisesti merkitsevä (Vaihe I: $\tau=.52$; Vaihe II: $\tau=.30$). Toisella hakukerralla oltiin kuitenkin taipuvaisempia hyväksymään aiheenmukaiseksi

runsaammin myös osittain tai mahdollisesti hyödyllisiä viitteitä. Molemmilla hakukerroilla kaikki osittain ja mahdollisesti hyödyllisiksi katsotut viitteet oli kuitenkin arvioitu vähintään heikosti aiheenmukaisiksi ja täysin hyödyllisiksi arvioidut vähintään jossain määrin aiheenmukaisiksi. Vaikuttaisikin siltä, että hyödyllisyys ja aiheenmukaisuus merkitsevät käyttäjille lähes samaa. Tehtävän kannalta hyödyllisiksi valitut viitteet olisivat siis aina myös aiheenmukaisia, mikä on varsin loogista, sillä tehtävän kannalta hyödyllisten viitteiden voi olettaa aina liittyvän myös valittuun aiheeseen. Tämä ei kuitenkaan sulje pois mahdollisuutta, etteikö käyttäjä saattaisi valita hyödylliseksi täysin epäaiheenmukaisia viitteitä, mutta tällöin näiden viitteiden voisi vastaavasti olettaa olevan käyttäjän mielestä hyödyllisiä jonkin muun tehtävän suhteen, kuin mitä varten tiedonhaku oltiin alun perin tekemässä. Tällöin käyttäjä tosin ei tule erotelleeksi aiheenmukaisuutta kahden eri tehtävän välillä. Tulosten perusteella psykologian opiskelijat olivat kuitenkin selvästi orientoituneet ainoastaan tutkimussuunnitelman laadintaan liittyvän tiedon hankkimiseen.

Viitteet todettiin sitä hyödyllisemmiksi mitä osuvammin ne tarjosivat yleistä taustatietoa, yksityiskohtaista tietoa, vinkkejä ongelman muotoilemiseen ja tuloksia tukemaan aiheenvalintaa. Ongelman muotoilussa auttavien vinkkien merkitys näyttäisi korostuneen toisella kerralla enemmän kuin viitteiden hyödyllisyys (Vaihe I: $\tau=,25$; Vaihe II: $\tau=,42$). Viitteiden hyödyllisyydestä ei oltu puolestaan niin yksimielisiä menetelmätiedon, empiriatiedon ja aiheen valintaa tukevien ideoiden osalta. Empiria- ja menetelmätietoa sisältäneet viitteet arvioitiin tehtävän loppupuolella osittain ja mahdollisesti hyödyllisiksi ehkä siksi, että näitä tietotyyppisiä ei arvioitu muutenkaan löydetyn kovin osuvasti. Löytyneen empiria- ja menetelmätiedon hyödyllisyydestä ei välttämättä oltu myöskään kovin varmoja.

6.5. Viitteiden aiheenmukaisuuden piirteet

Ensimmäisellä hakukerralla löytyneistä viitteistä suurin osa (46%) arvioitiin osuvasti aiheenmukaisiksi ja lähes yhtä paljon jossain määrin aiheenmukaisiksi (44%). Heikosti aiheenmukaisia oli vähiten (9%). Toisella hakukerralla oli eniten jossain määrin aiheenmukaisia (47%), osuvasti aiheenmukaisia seuraavaksi eniten (39%) ja vähiten heikosti aiheenmukaisia (14%). Kummallakaan kerralla ei valittu hyödylliseksi viitteitä, jotka olisi arvioitu ei lainkaan aiheenmukaisiksi.

Taulukko 7. Osovasti aiheenmukaisten viitteiden tietotyyppiarvioiden ja hyödyllisyysarvion keskiarvot (1= kuvaa osuvasti, 2= jossain määrin, 3= heikosti, 4= ei lainkaan), sekä erojen tilastollinen merkitsevyys.

Tietotyyppien keskiarvot	Osovasti aiheenmukaiset viitteet		
	Vaihe I	Vaihe II	p
Yleistä taustatietoa	1,5	1,8	.032
Tiettyä yksityiskohtaista tietoa	1,6	1,8	.571
Menetelmätietoa	2,1	2,3	.165
Empiirisiä tutkimustuloksia	1,8	2,0	.486
Ideoita aiheen valitsemiseen	2,2	3,0	.000
Vinkkejä ongelman muotoilemiseen	2,0	2,3	.116
Tuloksia tukemaan valittua aihetta	1,8	2,0	.365
Hyödyllisyyden keskiarvo	1,7	1,9	.204
<i>n</i>	70	38	

Aiheenmukaisuuden selittämiseksi tietotyyppien avulla on taulukossa 7 laskettu tietotyyppiarvioiden keskiarvot 'osuvasti aiheenmukaisiksi arvioituille' viitteille. Molemmilla hakukerroilla osuvasti aiheenmukaisiksi katsottujen viitteiden arvioitiin antavan työlle eniten yleistä taustatietoa (ka=1,5 ja 1,8) ja yksityiskohtaista tietoa (ka=1,6 ja 1,8) sekä vähiten ideoita aiheen valitsemiseen (ka=2,2 ja 3,0). Toisella hakukerralla kaikkien tietotyyppien arvioitiin kuvaavan vain hieman heikommin osuvasti aiheenmukaisia viitteitä, minkä vuoksi keskiarvoerot vaiheiden välillä olivat pienet. U-testin perusteella yleisen taustatiedon osalta muutos oli kuitenkin melkein merkitsevä (p=.032) ja aiheenvalintaa tukevien ideoiden osalta erittäin merkitsevä (p=.000).

Taulukko 8. Löytyneiden hyödyllisten viitteiden aiheenmukaisuuden ja tietotyyppiannin korrelaatiot eri hakukerroilla (**p<.010; *p<.050).

Tietotyyppi	Aiheenmukaisuus vaiheessa I		Aiheenmukaisuus vaiheessa II	
	/ viite	/ henkilö	/ viite	/ henkilö
Yleistä taustatietoa	,49**	,04	,38**	,57**
Yksityiskohtaista tietoa	,28**	,10	,22*	,28
Menetelmätietoa	,20**	,23	,17	,17
Empiirisiä tutkimustuloksia	,20**	,07	,09	,22
Ideoita aiheen valitsemiseen	,21**	,17	,07	-,01
Vinkkejä ongelman muotoilemiseen	,25**	,15	,29**	,21
Tuloksia tukemaan aiheenvalintaa	,42**	,20	,42**	,52**
<i>n</i>	150	21	98	19

Taulukossa 8 on esitetty viitteille ja henkilöille laskettu aiheenmukaisuuden ja tietotyyppien välinen korrelaatio. Viitteittäin mitattuna aiheenmukaisuus korreloi ensimmäisellä hakukerralla merkitsevästi kaikkien tietotyyppien kanssa, mutta kuitenkin eniten taustatiedon (tau=,49) ja aiheenvalintaa tukevien tulosten (tau=,42) kanssa. Henkilöittäin mitattuna kuitenkin ilmenee, ettei mitään tietyn tyyppistä tietoa todettu erityisen aiheenmukaiseksi. Käyttäjien mielestä menetelmä-

tieto kuitenkin korreloi eniten aiheen kanssa ($\tau=.23$), mikä on ymmärrettävää, sillä psykologian opiskelijoiden tehtävänä oli nimenomaan laatia suunnitelma menetelmiin perustuvaa tutkimusta varten. Toisella hakukerralla aiheenmukaisuus ($\tau=.42$) korreloi viitteittäin mitattuna merkitsevästi aiheenvalintaa tukevien tulosten, taustatiedon ($\tau=.38$) ja ongelman muotoilussa auttavien vinkkien ($\tau=.29$) kanssa sekä melkein merkitsevästi yksityiskohtaisen tiedon ($\tau=.22$) kanssa. Henkilöittäin mitattuna käyttäjät pitivät aiheenmukaisimpina lähinnä viitteitä, joiden arveltiin antavan yleistä taustatietoa ($\tau=.57$) ja tuloksia tukemaan valittua aihetta ($\tau=.52$).

Molemmilla hakukerralla katsottiin aiheenmukaisimmiksi aiheenvalintaa tukevat tulokset ja yleinen taustatieto. Myös molemmilla kerroilla, mutta erityisesti toisella kerralla, aiheenmukaisuus korreloi vähiten aiheenvalintaa tukevien ideoiden, menetelmätiedon ja empiriatiedon kanssa. Tietotyyppien heikko korrelaatio aiheenmukaisuuden kanssa saattaa viitata näiden tietotyyppien olevan sellaisia, joita käyttäjät etsivät myös oman spesifin aihealueen ulkopuolelta.

6.6. Haun onnistuneisuus yleensä

Psykologian opiskelijoilta kysyttiin tiedonhakujen lopuksi kokonaisvaltaisempaa käsitystä haun onnistuneisuudesta. Käyttäjää pyydettiin arvioimaan, kuinka kattavasti he uskoivat löytäneensä uusia viitteitä tehtäväänsä varten, ja miten tyytyväisiä he olivat saamiinsa viitteisiin kokonaisuudessaan.

Taulukko 9. Haun kattavuusarvio (1= hyvin kattava; 5= ei lainkaan kattava) ja tyytyväisyys saatuihin viitteisiin kokonaisuudessaan (1= hyvin tyytyväinen, 2=melko tyytyväinen, 3=melko tyytymätön, 4= täysin tyytymätön). Keskiarvot per henkilö hakukerroittain sekä erojen tilastollinen merkitsevyys vaiheitten välillä.

Haun onnistuneisuus	Vaihe I	Vaihe II	p
Haun arvioitu kattavuus	3,1	3,6	.211
Tyytyväisyys viitteisiin kokonaisuudessaan	2,2	2,4	.438
<i>n</i>	22	22	

Käyttäjien käsitys haun kattavuudesta mitattiin lomakkeessa viisiportaisella mittajanalla, johon oli luokiteltu ainoastaan muuttujan ääripäät (1= haku oli hyvin kattava, 5= haku ei ollut lainkaan kattava). Käyttäjien keskimääräistä käsitystä haun kattavuudesta ei voida tarkoin nimetä, mutta keskiarvojen perusteella haut arvioitiin toisella kertaa ($\bar{x}=3,6$) jonkin verran vähemmän kattaviksi kuin ensimmäisellä kerralla ($\bar{x}=3,1$) (Taulukko 9). Tehtävän alussa käyttäjillä ei näyttänyt olevan erityisiä odotuksia hakujen kattavuuden suhteen, sillä keskiarvo asettuu melko tarkoin janan keski-

vaiheille. Sen sijaan tehtävän lopussa käyttäjät olivat jonkin verran kriittisempiä arvioidessaan hakujen kattavuutta. Aihetietämyksen lisääntyessä käyttäjät mahdollisesti pystyivät paremmin arvioimaan, milloin omaan aiheeseen liittyvää tietoa oli löydetty kattavasti.

Tyytyväisyyttä haun tuloksena saatuihin viitteisiin kokonaisuudessaan mitattiin neliluokkaisella asteikolla (1= hyvin tyytyväinen saatuihin viitteisiin kokonaisuudessaan, 2=melko tyytyväinen, 3=melko tyytymätön, 4= täysin tyytymätön). Käyttäjät olivat keskimäärin tyytymättömämpiä viitteisiin kokonaisuudessaan toisella hakukerralla (ka=2,4) kuin ensimmäisellä kerralla (ka=2,2). Tehtävän alkupuolella käyttäjät olivat siis keskiarvon perusteella melko tyytyväisiä saamiinsa viitteisiin. Yksityiskohtaiseksi arvioidun tiedon löytyminen kompensoi mahdollisesti yleisen taustatiedon vähäistä löytymistä. Tehtävän lopussa käyttäjät ilmeisesti osasivat paremmin suhteuttaa odotuksensa tietokannan avulla löytyvien viitteitten antiin, sillä tyytyväisyysarvio asettui melko tarkoin mitta-asteikon keskivaiheille.

Taulukko 10. Löytyneiden hyödyllisten viitteiden piirteiden korrelaatiot suhteessa viitteiden määrään ja haun onnistuneisuusarvioihin laskettuna henkilöittäin (* $p \leq .01$; * $p \leq .05$) (ka. = keskiarvo per henkilö, katt.= kattavuus, tyyt.= tyytyväisyys).

Löytyneiden hyödyllisten viitteiden piirteet	Vaihe I			VaiheII		
	Määrä	Katt.	Tyyt.	Määrä	Katt.	Tyyt.
Määrä		-,50**	-,66**		-,56**	-,47*
Kattavuus			,73**			,70**
Täysin hyödyllisten osuus	,21	-,59**	-,45*	,51**	-,67**	-,57**
Aiheenmukaisuus (ka.)	-,26	,70**	,47*	-,18	,35	,48*
Yleistä taustatietoa (ka.)	-,19	-,01	-,01	-,38*	,40*	,40*
Yksityiskohtaista tietoa (ka.)	,02	-,03	,15	,29	-,02	,00
Menetelmätietoa (ka.)	,03	,10	,20	,35	,03	,04
Empiirisiä tutkimustuloksia (ka.)	-,28	,07	,24	,16	,05	,19
Ideoita aiheen valitsemiseen (ka.)	-,04	,07	,10	-,35	,26	,29
Vinkkejä ongelman muotoilemiseen (ka.)	-,17	,17	,37	-,43*	,50**	,43*
Tuloksia tukemaan aiheenvalintaa (ka.)	-,15	,27	,24	-,14	,38*	,43*
<i>n</i>	21	21	21	19	19	19

Taulukossa 10 on laskettu löytyneiden hyödyllisten viitteiden määrän, haun kattavuusarvioiden ja saatujen viitteiden kokonaistyytyväisyyden väliset korrelaatiot molemmilla hakukerroilla. Lisäksi taulukossa on esitetty täysin hyödyllisten viitteiden osuuden, kaikkien löytyneiden hyödyllisten viitteiden aiheenmukaisuuden sekä viitteiden tietotyypin korrelaatiot suhteessa löytyneiden

hyödyllisten viitteiden määrään¹⁶ ja haun onnistuneisuusarvioihin henkilöittäin¹⁷. Mitä enemmän viitteitä käyttäjät löysivät ensimmäisellä hakukerralla sitä kattavampina he pitivät hakuja (tau=-,50) ja sitä tyytyväisempiä he olivat saamiinsa viitteisiin kokonaisuudessaan (tau=-,66). Saatuihin viitteisiin oltiin myös tyytyväisiä kokonaisuudessaan, kun haku koettiin kattavaksi (tau=,73). Mitä suurempi osuus hyödyllisistä viitteistä oli puolestaan täysin hyödyllisiä sitä kattavamiksi haut koettiin (tau=-,59) ja sitä tyytyväisempiä oltiin myös saatuihin viitteisiin kokonaisuudessaan (tau=-,45). Aiheenmukaisuuden ja kattavuuden välinen korrelaatio oli erittäin merkitsevä (tau=-,70), mutta saatuihin viitteisiin oltiin myös tyytyväisiä, kun viitteet olivat aiheenmukaisia (tau=-,47). Tulosten perusteella käyttäjät kokivat haut kattaviksi ja olivat siis tyytyväisiä saamiinsa viitteisiin, kun viitteiden arveltiin olevan erityisesti täysin hyödyllisiä ja kun viitteet olivat aiheenmukaisia. Viitteiden tietotyyppiannin ja haun onnistuneisuuden mittareiden heikot korrelaatiot ilmaisevat puolestaan sen, että käyttäjät eivät ensimmäisellä kerralla löytäneet systemaattisesti mitään erityistä tiedon tyyppiä kattavasti tai tyydyttävästi.

Toisella hakukerralla käyttäjät pitivät hakuja kattavina löytäessään viitteitä paljon (tau=,56), mutta hakuihin ei sen sijaan oltu enää yhtä tyytyväisiä viitteiden määrän perusteella (tau=,47) kuin ensimmäisellä kerralla. Hakijat olivat kuitenkin tyytyväisiä saamiinsa viitteisiin, kun niitä löydettiin kattavasti (tau=,70). Käyttäjät, jotka olivat taipuvaisia arvioimaan viitteistä suuren osan täysin hyödyllisiksi, löysivät myös määrällisesti paljon viitteitä (tau=,51). Mitä suurempi osa viitteistä oli täysin hyödyllisiä sitä kattavampana myös pidettiin hakuja (tau=-,67) ja sitä tyytyväisempiä oltiin myös saatuihin viitteisiin kokonaisuudessaan (tau=-,57). Aiheenmukaisuus ei ollut toisella kerralla yhtä tärkeää kattavuuden kannalta kuin tehtävän alussa (tau=,35), mutta mitä aiheenmukaisempia viitteet kuitenkin olivat sitä tyytyväisempiä niihin myös oltiin kokonaisuudessaan (tau=,48). Osa käyttäjistä löysi kattavasti yleistä taustatietoa (tau=,40), vinkkejä ongelman muotoiluun (tau=,50) ja tuloksia tukemaan aiheen valintaa (tau=,38). Verrattaessa tätä taulukossa 3 esitettyyn faktorointiin, näyttääkin siltä, että toisella kerralla systemaattisemmin haussa kokivat onnistuneensa lähinnä ne käyttäjät, jotka löysivät aihetta tukevaa ja ongelman muotoilussa auttavaa yleistä tietoa.

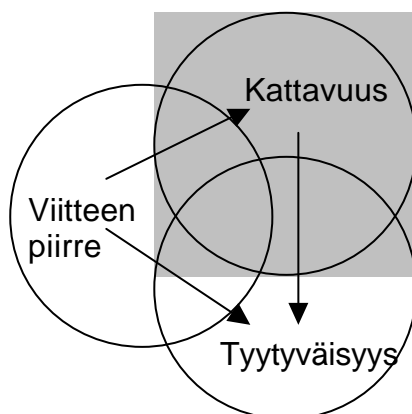
¹⁶ **Huomautus:** Kaikkien löytyneiden hyödyllisten viitteiden määrän ja täysin hyödyllisten viitteiden osuuden asteikot ovat erisuuntaiset suhteessa muihin muuttujiin.

¹⁷ Kullekin henkilölle on laskettu tämän löytämien viitteiden hyödyllisyys-, aiheenmukaisuus- ja tietotyyppiarvioiden keskiarvot.

Taulukko 11. Osittaiskorrelaatiot viitteiden piirteiden (v) ja saatujen viitteiden kokonaistyytyväisyyden (t) välillä, kun arvio haun kattavuudesta on vakioitu (τ_{vt} = vakioimaton korrelaatio taulukosta 10, $\tau_{vt,k}$ = kattavuuden suhteen vakioitu korrelaatio) (** $p \leq .01$; * $p \leq .05$).

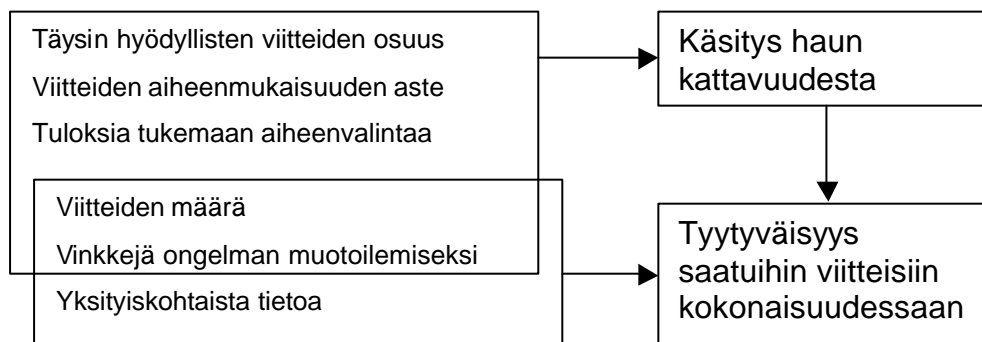
Löytyneiden hyödyllisten viitteiden piirteet (v)	Tyytyväisyys saatuihin viitteisiin vaiheessa I		Tyytyväisyys saatuihin viitteisiin vaiheessa II	
	τ_{vt}	$\tau_{vt,k}$	τ_{vt}	$\tau_{vt,k}$
Määrä	-,66**	-,50*	-,47*	-,13
Täysin hyödyllisten osuus	-,45*	-,04	-,57**	-,19
Aiheenmukaisuus (ka.)	,47*	-,08	,48*	,35
Yleistä taustatietoa (ka.)	-,01	,00	,40*	,18
Yksityiskohtaista tietoa (ka.)	,15	,25	,00	,02
Menetelmätietoa (ka.)	,20	,19	,04	,03
Empiirisiä tutkimustuloksia (ka.)	,24	,28	,19	,22
Ideota aiheen valitsemiseen (ka.)	,10	,07	,29	,16
Vinkkejä ongelman muotoil. (ka.)	,37	,37	,43*	,13
Tuloksia tukemaan aiheenvalint. (ka.)	,24	,07	,43*	,25
<i>n</i>	21	21	19	19

Taulukkoa 10 on käytetty edelleen apuna osittaiskorrelaatioiden laskemiseen, kun pyritään selvittämään, mistä kattavuus ja tyytyväisyys saatuihin viitteisiin mahdollisesti muodostuu. Taulukossa 11 on laskettu viitteiden määrän ja viitteiden eri piirteiden osittaiskorrelaatiot suhteessa saatujen viitteiden kokonaistyytyväisyyteen, kun kattavuuden vaikutus on vakioitu. Osittaiskorrelaatiota tehtäessä oletettiin tyytyväisyyden olevan riippuvainen kattavuudesta, sillä tulosten perusteella käyttäjät olivat aina myös tyytyväisiä saamiinsa viitteisiin, jos he pitivät hakuja kattavina. Sen sijaan saatuihin viitteisiin voitiin olla tyytyväisiä, vaikka hakua olisikaan pidetty kattavana. Osittaiskorrelaation perusteella saatiin selville, milloin käyttäjät olivat tyytyväisiä johonkin viitteen piirteeseen riippumatta kattavuudesta ja milloin piirteet olivat puolestaan sidoksissa kattavuusarvioihin (Kuvio 8).



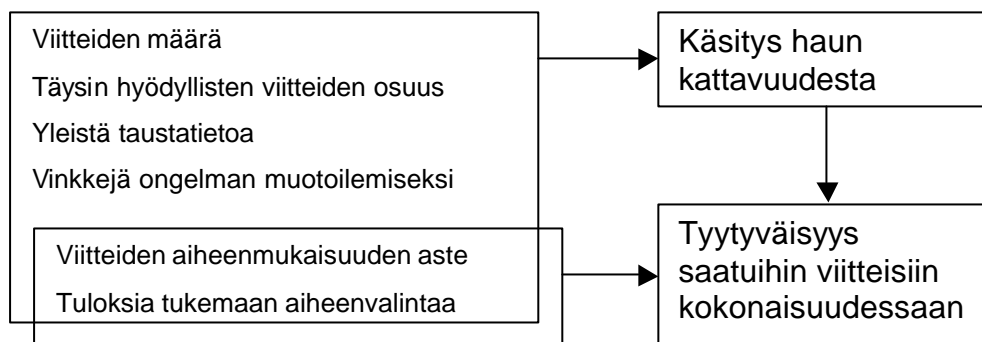
Kuvio 8. Saatujen viitteiden kokonaistyytyväisyyden selittyminen viitteen piirteiden perusteella, kun kattavuus on vakioitu.

Ensimmäisellä hakukerralla käyttäjät olivat kattavuudesta riippumatta tyytyväisiä saamiinsa viitteisiin, kun löysivät määrällisesti paljon viitteitä ($\tau_{vt,k}=-,50$) (Taulukko 11). Lisäksi käyttäjät näyttivät olleen tyytyväisiä saamiinsa viitteisiin löytäessään ongelman muotoilussa auttavia vinkkejä ($\tau_{vt,k}=,37$) tai yksityiskohtaista tietoa ($\tau_{vt,k}=,25$), vaikka he eivät olisikaan pitäneet hakua välttämättä kattavana. Viitteiden määrä ja ongelman muotoilussa auttavien vinkkien löytyminen korreloivat kuitenkin myös kattavuuden kanssa. Sen sijaan käyttäjät olivat tyytyväisiä saamiinsa viitteisiin täysin hyödyllisten viitteiden osuuden ($\tau_{vt,k}=-,04$), aiheenmukaisuuden ($\tau_{vt,k}=-,08$) ja aiheenvalintaa tukevien tulosten ($\tau_{vt,k}=,07$) perusteella lähinnä vain silloin, kun pitivät hakua myös kattavana. Ensimmäisellä hakukerralla täysin hyödyllisten osuus, aiheenmukaisuus ja aiheenvalintaa tukevien tulosten löytyminen vaikuttivat siis käsitykseen kattavuudesta ja sitä kautta tyytyväisyyteen. Viitteiden määrän ja piirteiden sekä haun onnistuneisuusarvioiden riippuvuudet ensimmäisellä hakukerralla on havainnollistettu tarkemmin kuviossa 9.



Kuvio 9. Haun onnistuneisuusarvioiden keskinäinen riippuvuus sekä riippuvuus viitteiden tietotyyppiannista, aiheenmukaisuudesta, määrästä ja täysin hyödyllisten viitteiden osuudesta ensimmäisellä hakukerralla.

Toisella kerralla käyttäjät olivat tyytyväisiä saamiinsa viitteisiin aiheenmukaisuuden ($\tau_{vt,k}=,35$) ja aiheenvalintaa tukevien tulosten ($\tau_{vt,k}=,25$) perusteella kattavuudesta riippumatta. Viitteiden määrän ($\tau_{vt,k}=-,13$), täysin hyödyllisten viitteiden osuuden ($\tau_{vt,k}=-,19$), yleisen taustatiedon ($\tau_{vt,k}=,18$) ja ongelman muotoilussa auttavien vinkkien ($\tau_{vt,k}=,13$) perusteella viitteisiin oltiin tyytyväisiä yleensä vain, kun hakua pidettiin myös kattavana. Viitteiden määrän ja piirteiden sekä haun onnistuneisuusarvioiden riippuvuudet toisella hakukerralla on havainnollistettu tarkemmin kuviossa 10.



Kuvio 10. Haun onnistuneisuusarvioiden keskinäinen riippuvuus sekä riippuvuus viitteiden tietotyyppiannista, aiheenmukaisuudesta, määrästä ja täysin hyödyllisten viitteiden osuudesta toisella hakukerralla.

Hakujen onnistuneisuutta voidaan tarkastella myös tietotyyppiä odotusten ja viitteiden tietotyyppiannin välisen eron perusteella. Taulukkoa 12 varten on ensin laskettu kullekin henkilölle tämän löytämien hyödyllisten viitteiden tietotyyppiä odotusten keskiarvo. Tämän jälkeen kunkin tietotyyppiä odotusten arvosta vähennetty saannin arvo. Tällöin negatiivinen arvo ilmaisee, että henkilön löytämät viitteet eivät ole antaneet kyseisen tyyppistä tietoa odotuksia vastaavasti ja päinvastoin. Näin saatu asteikko on korreloitu kattavuus- ja tyytyväisyysarvioiden kanssa. Koska asteikot ovat vastakkaiset, negatiivinen arvo ilmaisee tällöin, että kyseistä tietotyyppiä on löydetty odotettua enemmän ja että tämä on mahdollisesti voimistanut käsitystä haun kattavuudesta tai lisännyt tyytyväisyyttä saatuihin viitteisiin.

Taulukko 12. Haun onnistuneisuusarvioiden korrelaatiot suhteessa tietotyyppiä odotusten ja viitteiden tietotyyppiannin väliseen eroon (* $p \leq .05$).

Tietotyyppiä odotusten ja viitteiden tietotyyppiannin (ka/henk.) ero	Vaihe I		Vaihe II	
	Kattav.	Tyytyv.	Kattav.	Tyytyv.
Yleistä taustatietoa	,12	,02	-,14	-,30
Yksityiskohtaista tietoa	,10	-,08	-,03	-,09
Menetelmätietoa	,14	-,06	,01	-,06
Empiirisiä tutkimustuloksia	-,10	-,39*	,03	-,02
Ideoita aiheen valitsemiseen	-,28	-,05	-,29	-,31
Vinkkejä ongelman muotoilemiseen	,06	,05	-,08	-,27
Tuloksia tukemaan aiheenvalintaa	-,32	-,38	-,18	-,19
<i>n</i>	21	21	19	19

Tulosten perusteella käyttäjät olivat ensimmäisellä hakukerralla tyytyväisiä saamiinsa viitteisiin, mikäli he löysivät odotettua enemmän empiirisiä tutkimustuloksia ($\tau = -,39$) ja aiheenvalintaa tukevia tuloksia ($\tau = -,38$). Käyttäjät olivat myös mahdollisesti tyytyväisiä haun kattavuuteen, kun olivat löytäneet odotettua enemmän ideoita aiheen valitsemiseen ($\tau = -,28$) ja aiheen valintaa

tukevia tuloksia ($\tau=-,32$). Toisella hakukerralla käyttäjät olivat mahdollisesti tyytyväisiä saamiinsa viitteisiin, mikäli olivat löytäneet odotettua enemmän yleistä taustatietoa ($\tau=-,30$) ja vinkkejä ongelman muotoilemiseen ($\tau=-,27$). Myös ideoita aiheen valitsemiseen löydettiin toisella kerralla odotettua enemmän ja tämä korreloi lievästi niin kattavuuden kuin tyytyväisyydenkin kanssa. Tämän ei kuitenkaan voi olettaa juurikaan vaikuttaneen arvioihin haun onnistuneisuudesta, koska aiheen valinnassa auttavien ideoiden katsottiin kuvaavan vain heikosti löytyneitä hyödyllisiä viitteitä.

7. Tulosten tarkastelu

Tehdyn tutkimuksen yhtenä tavoitteena oli laajentaa Hakalan ja Vakkarin keväällä 1999 tekemää tehtäväkeskeistä relevanssitutkimusta, jossa tarkasteltiin informaatiotutkimuksen opiskelijoiden tiedonhankintaa heidän laatiessaan tutkimussuunnitelmaa pro gradu –tutkielmaa varten. Informaatiotutkimuksen opiskelijat olivat ensinnäkin opintojensa loppuvaiheessa ja toiseksi heitä voi pitää keskimääräistä kokeneempina tiedonhakijoina. Tässä tutkimuksessa tarkastellut psykologian opiskelijat puolestaan osallistuivat aineopintojen kurssille, jolloin he eivät välttämättä olleet yhtä perehtyneet oppialaansa kuin informaatiotutkimuksen opiskelijat. Psykologian opiskelijoiden hakutaidot ja hakukokemus olivat myös selvästi vähäisemmät.

Käyttäjäkeskeisiä tietotyypin odotusten, tietotyypin ja haun kokonaisuonnistuneisuuden välisiä suhteita ei sen sijaan ole aiemmin tutkittu. Lisäksi tutkimuksessa on esitetty muutamia uusia tietotyyppisiä sekä **mahdollisen hyödyllisyyden** -käsite. Tulosten perusteella psykologian opiskelijoiden löytämien hyödyllisten viitteiden tietotyypin poikkesi selvästi tietotyypin odotuksista erityisesti tehtävän alussa. Opiskelijat löysivät selvästi yksityiskohtaisempaa tietoa, kuin mitä he arvioivat tarvitsevansa ennen tiedonhakua. Tämä johtuu mahdollisesti opiskelijoiden alakohtaisesta kokemattomuudesta, minkä vuoksi suurella osalla käyttäjistä ei myöskään ollut kokemusta PsychINFO-tietokannan käytöstä tai annista. Tehtävän lopussa käyttäjät puolestaan arvioivat tarvitsevansa selvästi yksityiskohtaisempaa tietoa sekä menetelmä- ja empiriatietoa. Yksityiskohtaista tietoa löydettiin odotetussa määrin, mutta empiria- ja menetelmätietoa ei sen sijaan löydetty odotetusti. Opiskelijoiden tarvitsemat tutkimusmenetelmät saattoivat esimerkiksi olla niin spesifejä tai uusia, ettei niitä käsitteleviä artikkeleita löytynyt. On myös mahdollista, että käyttäjät eivät kummallakaan hakukerralla osanneet onnistuneesti muotoilla tehtävän kannalta keskeisiä käsitteitä hakutermeiksi ja hakulauseiksi. Hakutaitojen puuttuminen

saattoi kostautua erityisesti tehtävän lopussa, kun opiskelijat yrittivät löytää yksityiskohtaisia menetelmä- ja empiriatietoa käsitteleviä viitteitä.

7.1. Hyödyllisyysarviot eri käyttäjäryhmien välillä

Samoin kuin Vakkarin ja Hakalan (2000, 549) havainnoimat informaatiotutkimuksen opiskelijat, myös psykologian opiskelijat hyväksyivät tehtävän loppupuolella vähemmän viitteitä hyödyllisiksi kuin tehtävän alussa. Etenkin täysin hyödyllisiksi viitteitä hyväksyttiin tehtävän lopussa selvästi vähemmän. Tämä saattaa johtua aihetietämyksen myötä lisääntyneestä kriittisyydestä, jolloin käyttäjät pystyivät tehtävän lopussa arvioimaan paremmin dokumenttien hyötyä viitteiden perusteella (Vakkari 2000, 6). Mielestäni tiedonhauissa hyödyllisiksi hyväksytyjen viitteiden määrän vähenemiseen voi olla myös syynä tiedonhakujen tarpeen väheneminen tehtävän loppua kohden. Ensimmäisten tiedonhakujen avulla löytyneet viitteet saattavat esimerkiksi tarjota valmiita aihetta käsitteleviä viitteitä, jolloin uusien dokumenttien löytämiseksi ei tarvitse enää tehdä yhtä kattavia tiedonhakuja kuin tehtävän alussa. Lisäksi niin tässä kuin informaatiotutkimuksen opiskelijoillekin tehdystä tutkimuksesta osa käyttäjistä oli saattanut tehtävänsä ainakin siltä vaadittuun päätökseen, jolloin tiedon tarpeen voi olettaa olennaisesti vähentyneen. Toisaalta molemmissa tutkimuksissa opiskelijat siirtyivät tehtävästä suoriuduttuaan taas seuraavan, samaan aihealueeseen liittyvän, tehtävän pariin. Tällöin ainakin osa käyttäjistä katsoi edelleen tarvitsevansa uutta tietoa, mutta heidän aiheenrajauksensa oli muuttunut tehtävän seuraavaa vaihetta vastaavaksi.

Vakkarin ja Hakalan (2000, 550) tekemässä tutkimuksessa informaatiotutkimuksen opiskelijat löysivät tehtävän joka vaiheessa suhteellisesti enemmän täysin hyödyllisiä viitteitä kuin tässä tarkastellut psykologian opiskelijat. Etenkin tehtävän alussa suhde oli peräti kaksinkertainen. Informaatiotutkimuksen opiskelijat arvioivat ensimmäisellä hakukerralla 59% hyödyllisistä viitteistä täysin hyödyllisiksi, kun taas psykologian opiskelijat arvioivat täysin hyödyllisiksi 27% hyödyllisistä viitteistä. Tämä saattaa johtua joko informaatiotutkimuksen opiskelijoiden hakukokemuksesta tai yleisestä tieteeseen liittyvästä aihetietämyksestä. Informaatiotutkimuksen opiskelijat olivat kokeneempia hakijoita ja opinnoissaan pidemmällä kuin psykologian opiskelijat. Lisäksi informaatiotutkimuksen opiskelijat tunsivat oman tieteenalansa peruskäsitteet ja niiden väliset suhteet mahdollisesti paremmin. Hsieh-Yee (1993, 168) mukaan kokeneet hakijat yhdistävät tehokkaammin erityisesti synonyymeja hakulausekkeisiinsa, kun taas kokemattomat hakijat eivät taas tunne edes tarvetta synonyymien käyttöön. Tämän perusteella voidaan olettaa, että

informaatiotutkimuksen opiskelijat pystyivät muotoilemaan hakunsa paremmin, minkä vuoksi he löysivät vastaavasti myös asteeltaan hyödyllisempiä viitteitä.

7.2. Aiheenmukaisuuden ja hyödyllisyyden piirteet

Muun muassa Wang ja Soergel (1998, 128) ovat tutkimuksissaan esittäneet, että käyttäjät mahdollisesti valitsevat viitteitä tai dokumentteja riippumatta niiden aiheenmukaisuudesta, vaikka aiheenmukaisuus onkin yleensä tärkein valintakriteeri. Tässä tutkimuksessa saatujen tulosten perusteella näyttää kuitenkin siltä, että vaativaa tehtävää varten tehtyjen tiedonhakujen yhteydessä hyödyllisiksi katsottujen viitteiden oletetaan aina olevan ainakin heikosti aiheenmukaisia. Molemmilla hakukerroilla kaikki psykologian opiskelijoiden hyödyllisiksi valitsemat viitteet oli arvioitu vähintään heikosti aiheenmukaisiksi. Mielestäni aiheenmukaisuus onkin aina käyttäjien yhtenä kriteerinä relevanssiarvioita tehtäessä, vaikka käyttäjät eivät sitä erikseen mainitsisikaan.

Aiheenmukaisuuden tärkeys näyttäisi kuitenkin vähenevän tehtävän loppupuolella, sillä osuvasti aiheenmukaisten viitteiden osuus oli toisella hakukerralla pienempi kuin tehtävän alussa. Aihetuntemuksen lisääntyessä käyttäjät saattoivat valita hyödyllisiksi informaatio-objekteja, jotka eivät käsitelleet suoranaisesti heidän valitsemaansa rajattua aihetta. Informaatio-objektien oletettiin tällöin mahdollisesti tarjoavan tietoa, kuten esimerkiksi menetelmätietoa, joka on sovellettavissa omaan aiheeseen.

Tehtävän alussa arviot viitteiden aiheenmukaisuudesta, hyödyllisyyden asteesta ja tietotyyppien annista korreloivat voimakkaasti keskenään. Ennen oman aiheen rajauksen löytymistä käyttäjät pitivät viitteitä hyödyllisinä lähinnä vain silloin, kun ne tarjosivat myös osuvasti kaivatun tyyppistä tietoa ja kun viitteet olivat aiheenmukaisia. Tutkimuksessa käytetty hyödyllisyyden luokittelu puolestaan osoitti, että käyttäjät yleensä arvioivat täysin hyödyllisten katsottujen viitteiden tarjoavan enemmän odotetun tyyppistä tietoa kuin osittain hyödyllisten viitteiden. Osittain hyödyllisten viitteiden arvioitiin puolestaan yleensä tarjoavan osuvammin kaivatun tyyppistä tietoa kuin mahdollisesti hyödyllisten viitteiden. Tehtävän loppupuolella käyttäjät kuitenkin valitsivat menetelmä- ja empiriatietoa tarjoavia viitteitä usein hyödyllisyyden asteesta riippumatta. Oletettavasti lisääntyneen aihetietämyksen ansiosta käyttäjät havaitsivat viitteiden tarjoavan osuvasti kaivatun tyyppistä tietoa, vaikka käyttäjät eivät olleetkaan varmoja viitteiden hyödyllisyydestä. Tällöin he arvioivat viitteet yleensä **mahdollisesti aiheenmukaisiksi**.

7.3. Haun onnistuneisuus kaikkiaan

Käyttäjien arviot haun onnistuneisuudesta tukevat mielestäni vain osittain Sun (1994, 211) tekemää havaintoa, minkä mukaan käyttäjät asettavat informaatio-objektien laadun ja saannin etusijalle suhteessa hakujen tarkkuuteen. Lähinnä tehtävän alussa psykologian opiskelijat olivat suhteellisen tyytyväisiä saamiinsa viitteisiin ja pitivät hakuja kattavina, vaikka olisivat hakujen yhteydessä joutuneet käymään läpi suurenkin joukon viitteitä. Käyttäjät kuitenkin yrittivät haun aikana usein rajata löytämiensä hakujoukkojen kokoa. Monessa tapauksessa käyttäjä ei ryhtynyt edes selaamaan viitteitä ennen kuin hakujoukon koko oli onnistuttu rajaamaan muutamaa kymmeneen. Tämä näyttäisi osoittavan, että myös tarkkuus on hakuhetkellä tärkeä tekijä.

Tyytyväisyyttä saatuihin viitteisiin tarkasteltiin tässä tutkimuksessa haun kattavuudesta riippuvaisena tekijänä. Käyttäjät olivat siis aina myös tyytyväisiä saamiinsa viitteisiin, jos hakuja pidettiin kattavana. Sen sijaan käyttäjät saattoivat olla tyytyväisiä saamiinsa viitteisiin, vaikka hakua ei olisikaan pidetty kattavana. Kattavuuden ja tyytyväisyyden syy seuraussuhde ei kuitenkaan ole yksiselitteinen. Tulosten perusteella on kuitenkin selvää, että haun kattavuutta ja tyytyväisyyttä saatuihin viitteisiin arvioidaan niin määrän kuin myös laadullisten ominaisuuksien perusteella. Jossain määrin yllättävää kuitenkin on, että tehtävän alussa osa käyttäjistä näytti olevan tyytyväisiä saamiinsa viitteisiin määrän perusteella, vaikka eivät olisikaan pitäneet hakua kattavana. Ilmeisesti tehtävän alussa käyttäjät olivat tyytyväisiä, kun löysivät ylipäättään jotain mikä auttoi heitä pääsemään alkuun tehtävän suorittamisessa. Viitteitä ei sen sijaan uskottu löydettävän kattavasti ehkä siksi, että kattavuuden tiedostettiin olevan riippuvainen omasta aihetietämyksestä. Toisin sanoen oma tietämättömyyden tunne välittyi käsitykseen haun kattavuudesta. Tehtävän lopussa oltiin puolestaan tyytyväisiä saatuihin viitteisiin yleensä vain silloin, jos hakua pidettiin myös kattavana.

Muutamit käyttäjät olivat tehtävän lopussa erittäin tyytyväisiä saamiinsa viitteeseen eivätkä olleet erityisen pettyneitä haun kattavuuteenkaan, vaikka viitteitä oli löydetty vain pari kolme. Muutamissa tapauksissa käyttäjät olivat taas pettyneitä haun kattavuuteen selvästi viitteiden määrän takia. Tällöin pettymys heijastui myös tyytyväisyysarvioon. Eräissä tapauksissa käyttäjä oli esimerkiksi löytänyt vain yhden viitteen, jonka hän oli arvioinut tarjoavan osuvasti kaivatun tyyppistä tietoa. Käyttäjä ei ymmärrettävästi voinut pitää hakua sen enempiä kattavana kuin olla tyytyväinen saamiinsa viitteisiin kokonaisuudessaan, koska tyytyväisyyttä kysyttiin monikossa. Viitteiden määrä vaikutti toisella kerralla siis selvästi ensisijaisesti kattavuusarvioon ja sitä kautta myös tyytyväisyyteen. Tarkkuuden voi olettaakin olevan tärkeämpää tehtävän lopussa.

Tulosten perusteella näyttikin siltä, että psykologian opiskelijoille saanti oli tärkeää lähinnä tehtävän alussa. Tällöin käyttäjät pyrkivät selvästi löytämään paljon viitteitä laadusta riippumatta. Sen sijaan tehtävän lopussa oltiin kriittisempiä viitteiden laadun ja haun yleisen onnistuneisuuden suhteen. Tutkimuksessa olisikin kenties ollut syytä kysyä haun onnistuneisuutta myös suhteessa käytettyyn aikaan ja läpi käytyjen viitteiden määrään. Tällöin olisi ehkä käynyt selvemmin ilmi, mikä oli käyttäjien käsitys haun tarkkuudesta. Käyttäjät tosin saattavat olla tarkkuuden suhteen vähempäänkin tyytyväisiä, jos hakujärjestelmien ei ylipäätään odoteta pystyvän kovin suureen tarkkuuteen. Jos järjestelmän tarkkuutta voitaisiin parantaa erityisesti siten, että käyttäjät itse kokisivat löytävänsä kattavasti aiheenmukaista ja kaivatun tyyppistä tietoa, asettaisivat he todennäköisesti myös tarkkuudelle korkeammat kriteerit.

7.4. Tutkimuksen anti

Tutkimus ei mielestäni tarjoa juurikaan valmiita ratkaisuja hakujärjestelmien suunnittelulle. Tosin kattava tietotyyppien mukainen luokitus voisi ehkä antaa käyttäjille yhden uuden dokumenttien haettavuutta lisäävän keinon. Ingwersenin (1996, 25) mukaan kaikkien eri rooleissa toimivien käyttäjien tietämysten limittyminen (cognitive overlap) avaa myös yksittäiselle käyttäjälle enemmän mahdollisuuksia lähestyä aihetta omasta näkökulmasta. Tällöin tietotyyppijaottelu voisi toimia esimerkiksi asiasanoitusten, tesaaurusten tai luokitusten rinnalla yhtenä tiedon eri aspekteja kuvaavana piirteenä. Tietotyyppien mukainen luokittelu ei kuitenkaan ole ongelmatonta, sillä ainakin tässä tarkastellut tietotyypit ovat hyvin yksilöllisesti pääteltäviä. Esimerkiksi yhden käyttäjän mielestä yksityiskohtainen tieto voi olla toisen mielestä yleistä taustatietoa. Ideoita aiheen valintaan saatetaan puolestaan löytää lähes minkä tyyppisistä dokumenteista tahansa. Ingwersen (1996, 30) on tosin esittänyt, että dokumentteja voitaisiin tarkastella esimerkiksi tekstikappaleiden ja lukujen kokoisina semanttisina kokonaisuuksina. Haun voisi tällöin rajata dokumenttien niihin osioihin, joissa todennäköisimmin tietyn tyyppistä tietoa esiintyy. Tällöin haku tosin onnistuisi ainoastaan kokotekstitietokannoista.

Kirjastojen luokitusjärjestelmiä ja tesauroksia voidaan myös mielestäni pitää eräänlaisina tietotyyppijaotteluina. Luokitellut aihehakemistot saattavat olla kokemattomammillekin tiedonhakijoille tuttuja esimerkiksi Internetin aihehakupalvelujen yhteydestä. Aihehakemiston avulla tehtävän alussa oleva käyttäjä saattaa päästä tiedonhaussa ja -hankinnassa helpommin eteenpäin. Tiedonhaussa pidemmällä olevat saattaisivat puolestaan hyötyä tesaauruksen tarjoamista asiasanoista. Käyttäjät eivät mielestäni kuitenkaan ole kovin aktiivisia hakutermien valitsijoita.

Hakujärjestelmän pitäisikin tarjota käyttäjälle haun laajennusta automaattisesti. Järjestelmä esimerkiksi ehdottaisi käsitteellisesti suppeampia termejä käyttäjän valitsemille termeille. Järjestelmän ehdottamat termit voisivat tulla näkyville vaikkapa samalla tavoin kuin Internet-selaimien osoiteikkunan alle avautuvat käytyjen osoitteiden listat. Käyttäjää ei siis pakotettaisi ottamaan käyttöön järjestelmän ehdottamia termejä, mutta ne olisivat kuitenkin nopeasti saatavilla.

PsycInfossa ei ole käytettävissä aiheenmukaista hakemistoa, mutta viitteet on asiasanoitettu melko kattavasti ja käyttäjillä on mahdollisuus selata asiasanoja aakkosellisen indeksin tai tesauruksen avulla. Kokemattomalle käyttäjälle indeksin tai tesauruksen käyttö ei kuitenkaan ole kovin helppoa. Psykologian opiskelijoista vain muutama käytti indeksiä tai tesaurusta ja yleensä huonolla menestyksellä. Muutamat käyttäjät näyttivät olevan jopa hämmentyneitä, kun kirjoitusasullisesti täysin samalla asiasanalla löytyi vähemmän viitteitä kuin vapaatekstiksi määritellyllä hakutermillä. Tämänkaltaiset ongelmat voitaisiin mielestäni välttää selkeän ja opastavan käyttöliittymän avulla. Käyttäjän tulisi löytää ja tunnistaa apuväline, osata käyttää sitä ja ymmärrettävä myös sen toiminnan peruserä. Koska kaikkia viitteitä ja dokumentteja ei myöskään voida asiasanoittaa yksiselitteisesti tai luokitella kattavasti, pitäisi käyttäjälle myös ilmaista selkeästi, mitkä ovat asiasanoituksen ja luokituksen rajoitukset.

Käsitteellisesti tutkimus auttaa puolestaan selkiyttämään, mitä käyttäjakeskeinen relevanssi ja aiheenmukaisuus itse asiassa ovat. Mielestäni käyttäjakeskeisessä relevanssitutkimuksessa voitaisiin perustellusti käyttää termejä **hyödyllisyys** (utility) ja **hyöty** (use), kun tarkoitetaan viitteiden tai dokumenttien relevanttisuutta loppukäyttäjän näkökulmasta. Harterin (1992, 603) **psykologinen relevanssi** ja Saracevicin (1996, 214) **kognitiivinen relevanssi** vastaavat taas mielestäni aika tarkoin käyttäjän käsitystä aiheenmukaisuudesta. Käyttäjä voi siis tiedonhakuhetkellä arvioida informaatio-objektin aiheenmukaiseksi omien tuntemustensa ja käsityksensä perusteella. Informaatio-objektia ei kuitenkaan pidetä hyödyllisenä, jos sen ei uskota sisältävän kaivatun tyyppistä tietoa joka auttaa ongelman ratkaisussa ja tehtävän suorituksessa. Tämän tutkimuksen perusteella voidaan lisäksi olettaa, että ainakin vaativaa tehtävää suorittava käyttäjä valitsee hyödyllisiksi viitteitä, jotka ovat ainakin jollain tavoin myös aiheenmukaisia. Jos siis kaivatun tyyppistä tietoa löydetään, on sen myös väistämättä liityttävä ainakin osittain johonkin käyttäjän tehtävässään tarkastelemaan aiheeseen. Ongelmana voi esimerkiksi olla vaikkapa uuden tiedon omaksuminen tai vanhan tiedon tarkistaminen. Kummassakin tapauksessa tiedon tarve kuitenkin liittyy tehtävään ja tieto on siis käyttäjälle aiheenmukaista.

Lähteet

- Barry, C. (1994), User-Defined Relevance Criteria: an Exploratory Study. *JASIS* 45(3): 149-159.
- Barry, C. ja Schamber, L (1998), Users' Criteria for Relevance Evaluation: A Cross-situational Comparison. *Information Processing & Management* 34(2-3): 219-236.
- Bateman, J. (1998), Changes in Relevance Criteria: a Longitudinal Study. *Proceedings of the 61st ASIS Annual Meeting* 35:23-32.
- Bates, M. (1989), The Design of Browsing and Berry picking Techniques for the Online Search Interface. *Online Review* 13(5): 407-424.
- Cramer, D. (1994), *Introducing statistics for social research*. New York: Routledge.
- Engeström, Y. (1987), *Perustietoa opetuksesta*. Helsinki: Valtiovarainministeriö.
- Hakala, N. (1999), *Relevanssiarvioiden suhde ongelmanongelman jäsentymiseen. Arviointikriteerien ja tietotyypin muutokset tiedonhakuprosessissa*. Tampere: Pro gradu -työ Tampereen yliopisto informaatiotutkimuksen laitos.
- Harter, Stephen P. (1992), Psychological Relevance and Information science. *JASIS* 43(9): 602-615.
- Hert, C. (1996), User Goals on an Online Public Access Catalog. *JASIS* 47(7): 504-518.
- Hsieh-Yee, I. (1993), Effects of Search Experience and Subject Knowledge on the Search Tactics and Experienced Searches. *JASIS* 44(3): 161-174.
- Ingwersen, P. (1996), Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR Theory. *Journal of Documentation* 52(1): 3-50.
- Järvelin, K. (1995), *Tekstitiedonhaku tietokannoista. Johdatus periaatteisiin ja menetelmiin*. Espoo: Suomen ATK-kustannus.
- Kuhlthau, C. (1991), Inside the Search Process: Information Seeking from the Users' Perspective. *JASIS* 42(5): 361-371.
- Kuhlthau, C. (1993), *Seeking Meaning: A Process Approach to Library and Information Services*. Norwood, NJ: Ablex.
- Lowry, R. (1999-2000), *Partial correlation*. [viitattu 25.10.2001]. Saatavilla www-muodossa: <URL:<http://faculty.vassar.edu/lowry/ch3a.html>>.
- Metsämuuronen, J. (2000) *SPSS aloittelevan tutkijan käytössä*. Helsinki: International Methelp.
- Park, T. (1993), The Nature of Relevance in Information Retrieval: An Empirical Study. *Library Quarterly* 63(3): 318-351.
- Rautio, P. ja Suhonen, P. (1981), *Yhteiskuntatieteiden virrat ja suomalainen tutkija. Raportti Sosiologian ja valtio-opin väitöskirjojen lähteistä*. Helsinki: Suomen Akatemian julkaisu 2/1981.

- Park, T. (1993), The Nature of Relevance in Information Retrieval: An Empirical Study. *Library Quarterly* 63(3):318-351.
- Saracevic, T. (1975), Relevance: A Review of the Literature and Framework for Thinking on the Notion in Information Science. *JASIS* 26: 321-343.
- Saracevic, T. (1996a), Modelling Interaction in Information Retrieval (IR): a Review and Proposal. *Proceedings of the 59th Annual Meeting of the ASIS* 33: 3-9.
- Saracevic, T. (1996b), Relevance reconsidered '96. Teoksessa: Ingwersen, P. ja Pors, N.O. (toim.). *Information Science: Integration in Perspective*. Kööpenhamina: Royal School of Library and Information Science, 201-218.
- Savolainen, R. (1999), Tiedontarpeet ja tiedonhankinta. Teoksessa I Mäkinen (toim.) *Tiedon tie: Johdatus informaatiotutkimukseen*. Helsinki: BJT Kirjastopalvelu, 73-109.
- Schamber, L. (1991), User's Criteria for Evaluation in a Multimedia Environment. *Proceedings of the American Society for Information Science 54th Annual Meeting* 28: 126-133.
- Shannon, C. ja Weaver, W. (1963), *The Mathematical Theory of Communication*. Urbana (IL): University of Illinois Press.
- Spink, A., Greidorf, H. ja Bateman, J. (1998), From Highly Relevant to Non relevant: Examining Different Regions of Relevance. *Information Processing and Management* 34(5): 599-622.
- Su, L. (1994), The Relevance of Recall And Precision in User Evaluation. *JASIS* 45(3): 207-217.
- Vakkari, P. (1999), Task Complexity, Problem Structure and Information Actions: Integrating Studies on Information Seeking and Retrieval. *Information Prosessing and management* 35(1999): 819-837.
- Vakkari, P. ja Hakala, N. (2000), Changes in Relevance Criteria and Problem Stages in Task Performance. *Journal of Documentation* 56(5): 540-562.
- Vakkari, P. (2000), Relevance and Contributing Information Types of Searched Documents in Task Performance. *Proceedings of the American Society for Information Science Annual Meeting*.
- Vakkari, P. (2001), A Theory of the Task-Based Information Retrieval Process. *Journal of Documentation* 57(1): 44-60.
- Wang, P. ja Soergel D. (1998), A Cognitive Model of Document Use during a Research Project. Study I. Document Selection. *JASIS* 49(2): 115-133.

Liite 1: Process survey –lomake tiedonhankinnan vaiheen mittaamiseksi.

LOMAKE 1: PROCESS SURVEY

ID_____

1. Kuvaa aihepiiriäsi kirjallisesti muutamalla (5-10) lauseella ja piirrä käsitekartta aiheesi keskeisistä tekijöistä ja niiden välisistä suhteista.

2. Mikä on työsi nimi tällä hetkellä?

3. Missä määrin tunnet aiheitasi? Rengasta sopiva vaihtoehto.

-en lainkaan	1
-hieman	2
-kohtalaisesti	3
-runsaasti	4

4. Kenen kanssa olet puhunut projektistasi?

5. Arvioi itseluottamustasi tässä vaiheessa projektia asteikolla 1-10
(1=huono, 10=erinomainen)

6. Valitse seuraavista adjektiiveista kaksi sopivinta tärkeysjärjestyksessä
(1=tärkein, 2=seuraavaksi tärkein), jotka kuvaavat tuntemuksiasi projektin tässä vaiheessa.

- | | | | |
|---|---------------------------------------|--------------------------------------|--------------------------------------|
| <input type="checkbox"/> Luottavainen | <input type="checkbox"/> Hämmäntynyt | <input type="checkbox"/> Helpottunut | <input type="checkbox"/> Tyytyväinen |
| <input type="checkbox"/> Varma | <input type="checkbox"/> Epävarma | <input type="checkbox"/> Pettynyt | <input type="checkbox"/> Epäilevä |
| <input type="checkbox"/> Turhautunut | <input type="checkbox"/> Optimistinen | | |
| <input type="checkbox"/> Muu, mikä? _____ | | | |

7. Mikä on tehtäväsi tällä hetkellä? Merkitse kaksi sopivinta tärkeysjärjestykseen 1=tärkein, 2=seuraavaksi tärkein).

- Tehtävän alustava ymmärtäminen
- Työni yleisen aihepiirin identifiointi
- Työni yleiseen aihepiiriin liittyvän tiedon tutkiminen
- Tarkan aiheen muotoileminen
- Tiettyyn, tarkkaan aiheeseen liittyvän tiedon kerääminen
- Esitelmän viimeistely
- Muu, mikä? _____

8. Mitä teet tällä hetkellä esitelmääsi varten? Rastita kaikki sopivat vaihtoehdot.

- Keskustelen mahdollisista aiheista
- Teen kattavia tiedonhakuja
- Selailen tiedonlähteitä
- Luonnostelen työni jäsenystä
- Luen muistiinpanoja valitsemastani työn aiheesta
- Teen alustavia tiedonhakuja
- Keskustelen omaa aihettani tuntevien ihmisten kanssa
- Teen viimeisiä tiedonhakuja
- Kirjoitan muistiin valitsemani aihetta koskevia ideoita
- Luen valitsemaani aihetta koskevaa kirjallisuutta
- Teen yksityiskohtaisia muistiinpanoja faktoista ja ideoista
- Teen yleisluontoisia muistiinpanoja faktoista ja ideoista
- Tarkastan uudelleen lähteitä, jotka alunperin hylkäsin
- Kirjoitan muistiin julkaisujen viitteitä
- Muuta, mitä? _____

9. Mitä ajattelet parhaillaan tehtävästäsi? Rastita kaikki sopivat vaihtoehdot.

- Jäsenän tiedonlähteistä saamiani ideoita ja tietoja
- Mietin mahdollisia vaihtoehtoisia aiheita
- Tutustun aihepiiriin yleisesti
- Käyn läpi kaikki mahdolliset lähteet
- Harkitsen vaihtoehtoisia aiheita käytössäni olevien tietojen perusteella
- Olen valitsemassa toteuttamiskelpoisen aiheen
- Yritän alustavasti ymmärtää edessä olevan tehtävän
- Mietin tapoja viimeistellä tutkimussuunnitelmani
- Harkitsen vaihtoehtoisia aiheita käytettävissä olevan ajan suhteen
- Keskityn tiettyihin teemoihin työni yleisestä aihepiiristä
- Harkitsen vaihtoehtoisia aiheita tehtävän vaatimusten suhteen
- Selvittelen löytämieni tietojen ristiriitaisuuksia ja yhteensopimattomuuksia
- Kiinnostun aiheestani enemmän
- Jäsenän yksityiskohtaisesti tarkkaa aihettani
- Saavutan selvyuden ja suunnan tunteen
- Muistelen edellistä projektia, jossa etsin tietoa
- Pohdin, miten vaihtoehtoiset aiheen täsmennykset toimisivat
- Identifioin useita mahdollisia keskittymiskohteita työni yleisestä aihepiiristä
- Harkitsen vaihtoehtoisia aiheita omien henkilökohtaisen kiinnostusteni perusteella
- Hankin tietoa työni täsmennyksestä aiheesta
- Muuta, mitä? _____

Liite 2: Lomake hakukokemuksen testaamiseksi ensimmäisellä hakukerralla.

LOMAKE 2. Hakukokemuksen testaus

ID_____

1. Kuinka usein olet käyttänyt kuluneen 12 kk:n aikana kirjaston TAMCAT-näyttöluetteloa?

- 1 useita kertoja viikossa
- 2 kerran pari viikossa
- 3 joitakin kertoja kuukaudessa
- 4 muutaman kerran vuodessa
- 5 en lainkaan

2. Oletko käyttänyt seuraavia TAMCAT:n piirteitä?

Asiasanahaku (s/)

- 1 usein
- 2 harvoin
- 3 en lainkaan
- 4 en tunne hakukomentoa

Vapaasanahaku (s/)

- 1 usein
- 2 harvoin
- 3 en lainkaan
- 4 en tunne hakukomentoa

Luokitushaku (s/)

- 1 usein
- 2 harvoin
- 3 en lainkaan
- 4 en tunne hakukomentoa

Boolean haku (s/)

- 1 usein
- 2 harvoin
- 3 en lainkaan
- 4 en tunne hakukomentoa

3. Oletko osallistunut Tampereen yliopiston kirjaston tarjoamaan kirjastonkäytön opetukseen?

- 1 Kyllä, _____kertaa. Mihin kursseihin? _____
- 2 en

4. Kuinka usein olet käyttänyt PsychLit-tietokantaa?

- 1 en tunne järjestelmää
- 2 tunnen, mutta en ole käyttänyt
- 3 olen aloittanut käytön A9-kurssin yhteydessä
- 4 olen käyttänyt satunnaisesti ennen A9-kurssia
- 5 olen käyttänyt usein ennen A9-kurssia

Jos olet käyttänyt PsychLit/psychINFO -tietokantaa niin vastaa vielä seuraavaan kysymykseen. Muutoin siirry kysymykseen 5.

4a. Minkälaiseksi arvioit itse valmiutesi/kokemuksesi PsychINFO-tietokannan käytössä? Merkkää oma arviosi seuraavalle asteikolle.

|_____||_____||_____||_____||_____||

1
Aloitteleva käyttäjä
(ei juurikaan kokemusta)

5
Kokenut käyttäjä
(käytän järjestelmää usein)

5. Haet tietoa aikuisten television katsomisen ja aggression välisestä yhteydestä. Sopivia hakusanoja ovat mm.

aggression	television wieving	adults
aggressive behaviour	media	children
violence		

Havainnollista lyhyesti, miten yhdistäisit hakutermejä toisiinsa Boolean operaattoreilla ”AND, OR, NOT”. Kirjoita vastauksesi edellisen sivun kääntöpuolelle.

Liite 3: Lomake tiedon odotetun tarpeen mittaamiseksi tietotyypeittäin ensimmäisellä hakukerralla.

LOMAKE 3A: HAETUN INFORMAATION TIETOTYYPIT

ID_____

Miten seuraavat tietotyypit kuvaavat sitä, mitä haluat löytyvän kirjallisuuden tässä vaiheessa antavan työllesi? Rengasta sopiva vaihtoehto.

	Osuvasti	Jossain määrin	Heikosti	Ei lainkaan
Aihepiiriä koskevaa yleistä taustatietoa	1	2	3	4
Teorioita, malleja ja käsitelmäärittelyjä	1	2	3	4
Tietoa tutkimusmenetelmistä	1	2	3	4
Yksityiskohtaisia empiirisiä tutkimustuloksia	1	2	3	4
Ideoita aiheen valitsemiseksi	1	2	3	4
Vinkkejä kysymyksenasettelun täsmentämiseksi	1	2	3	4
Tiettyjä yksityiskohtaisia tietoja	1	2	3	4
Tuloksia tukemaan aihevalintaani	1	2	3	4

Liite 4: Lomake hakukokemuksen ja tietotyypiodotusten testaamiseksi toisella hakukerralla

LOMAKE 2. Hakukokemuksen testaus

ID _____

1. Kuinka monta kertaa olet käyttänyt PsychINFO/PsychLit-tietokantaa ensimmäisen haastattelukerran jälkeen? _____

2. Minkälaiseksi arvioit itse valmiutesi/kokemuksesi PsychINFO-tietokannan käytössä? Merkkää oma arviosi seuraavalle asteikolle.

1

Aloitteleva käyttäjä
(ei juurikaan kokemusta)

5

Kokenut käyttäjä
(käytän järjestelmää usein)

3. Miten seuraavat tietotyypit kuvaavat sitä, mitä haluat löytyvän kirjallisuuden tässä vaiheessa antavan työllesi? Rengasta sopiva vaihtoehto.

	Osovasti	Jossain määrin	Heikosti	Ei
lainkaan				
Aihepiiriä koskevaa yleistä taustatietoa	1	2	3	4
Teorioita, malleja ja käsittemäärittelyjä	1	2	3	4
Tietoa tutkimusmenetelmistä	1	2	3	4
Yksityiskohtaisia empiirisiä tutkimustuloksia	1	2	3	4
Ideoita aiheen valitsemiseksi	1	2	3	4
Vinkkejä kysymyksenasettelun täsmentämiseksi	1	2	3	4
Tiettyjä yksityiskohtaisia tietoja	1	2	3	4
Tuloksia tukemaan aihevalintaani	1	2	3	4

Liite 5: Viitteiden tietotyyppiantia mittaava lomake.

LOMAKE 3B: VIITTEIDEN TIETOTYYPIT ID _____/viiteno _____

Miten seuraavat seikat kuvaavat viitteen hyödyllisyyttä työllesi sen nykyvaiheessa? Rengasta sopiva vaihtoehto.

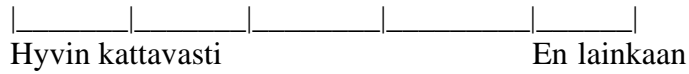
	Osuvasti	Jossain määrin	Heikosti	Ei
lainkaan				
On työni aihepiiristä	1	2	3	4
Sisältää aihepiiriä koskevaa yleistä taustatietoa	1	2	3	4
Antaa tietoa tutkimusmenetelmistä	1	2	3	4
Antaa yksityiskohtaisia empiirisiä tutkimustuloksia	1	2	3	4
Sisältää ideoita aiheen valitsemiseksi	1	2	3	4
Tuo vinkkejä kysymyksenasettelun täsmentämiseksi	1	2	3	4
Sisältää tiettyjä yksityiskohtaisia tietoja	1	2	3	4
Sisältää tuloksia tukemaan aihevalintaani	1	2	3	4
Tieto on tuoretta	1	2	3	4
Tekijä on arvostettu tutkija	1	2	3	4
Artikkeli on ilmestynyt arvostetussa julkaisussa	1	2	3	4

Liite 6: Lomake haun onnistuneisuuden mittaamiseksi.

LOMAKE 4: HAUN EVALUONTI

ID_____

1. Kuinka kattavasti uskot löytäneesi uusia viitteitä tehtävääsi varten? Merkkää arviosi seuraavalle jakaumalle.



2. Kuinka tyytyväinen ole saamiisi viitteisiin kokonaisuudessaan? Rengasta sopiva vaihtoehto.

- 1 hyvin tyytyväinen
- 2 melko tyytyväinen
- 3 melko tyytymätön
- 4 täysin tyytymätön

3. Löysitkö tiedonhaun antamien vihjeiden pohjalta uusia tehtävääsi koskevia käsitteitä?

- 1 kyllä
- 2 en

Mikäli löysit uusia käsitteitä työhösi, nimeä ne alle (mieluiten englanniksi).

	Ovatko ne (rengasta sopiva)			
	Täsmällisempiä	Synonyymejä	Laajempia	Täysin uusia
_____	1	2	3	4
_____	1	2	3	4
_____	1	2	3	4
_____	1	2	3	4
_____	1	2	3	4

4. Millaiseksi koit PsychINFO-järjestelmän käytön

	helpoksi	melko helpoksi	melko vaikeaksi	vaikeaksi tunne	en
Yleensä	1	2	3	4	5
kyselymuotoilun tuen käytön	1	2	3	4	5
tesauruksen käytön	1	2	3	4	5
indeksin käytön	1	2	3	4	5
kenttärajauksen käytön	1	2	3	4	5

Liite 7: Viitteiden tietotyyppien korrelaatiot (Kendallin tau_b)

Viitteiden tietotyyppien korrelaatiot vaiheessa I

		Yleistä	Yksit.	Menet.	Empir.	Ideoit.	Vinkk.	Tuloks.
Yleistä taustatietoa	tau_b	1,000						
	p	,						
	N	149						
Yksityiskoht. tietoa	tau_b	,184	1,000					
	p	,014	,					
	N	146	146					
Menetelmätietoa	tau_b	,136	,237	1,000				
	p	,066	,002	,				
	N	147	144	147				
Empiriatietoa	tau_b	,014	,605	,505	1,000			
	p	,853	,000	,000	,			
	N	146	143	144	146			
Ideoita aiheen valitsemiseen	tau_b	,378	,371	,233	,240	1,000		
	p	,000	,000	,001	,001	,		
	N	147	144	145	144	147		
Vinkkejä ongelman muotoilemiseen	tau_b	,341	,464	,191	,312	,697	1,000	
	p	,000	,000	,008	,000	,000	,	
	N	147	145	145	144	146	147	
Tuloksia tukemaan aiheenvalintaa	tau_b	,388	,530	,236	,378	,464	,506	1,000
	p	,000	,000	,001	,000	,000	,000	,
	N	147	144	145	144	146	146	147

Viitteiden tietotyyppien korrelaatiot vaiheessa II

		Yleistä	Yksit.	Menet.	Empir.	Ideoit.	Vinkk.	Tuloks.
Yleistä taustatietoa	tau_b	1,000						
	p	,						
	N	98						
Yksityiskoht. tietoa	tau_b	,021	1,000					
	p	,813	,					
	N	98	98					
Menetelmätietoa	tau_b	-,076	,119	1,000				
	p	,399	,178	,				
	N	98	98	98				
Empiriatietoa	tau_b	-,061	,494	,257	1,000			
	p	,502	,000	,004	,			
	N	97	97	97	97			
Ideoita aiheen valitsemiseen	tau_b	,208	,041	-,031	,073	1,000		
	p	,022	,645	,731	,417	,		
	N	98	98	98	97	98		
Vinkkejä ongelman muotoilemiseen	tau_b	,301	,302	,042	,136	,485	1,000	
	p	,001	,001	,627	,121	,000	,	
	N	98	98	98	97	98	98	
Tuloksia tukemaan aiheenvalintaa	tau_b	,314	,375	,234	,344	,260	,419	1,000
	p	,000	,000	,007	,000	,003	,000	,
	N	97	97	97	96	97	97	97