



# **KETRAMPILAN PENALARAN DEDUKTIF**

**(Deductive reasoning skills)**



YAYASAN PRIMA AGUS TEKNIK

**Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM**

# KETRAMPILAN PENALARAN DEDUKTIF

(Deductive reasoning skills)

**Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM**

## BIO DATA PENULIS



Penulis memiliki berbagai disiplin ilmu yang diperoleh dari Universitas Diponegoro (UNDIP) Semarang. dan dari Universitas Kristen Satya Wacana (UKSW) Salatiga. Disiplin ilmu itu antara lain teknik elektro, komputer, manajemen dan ilmu sosiologi. Penulis memiliki pengalaman kerja pada industri elektronik dan sertifikasi keahlian dalam bidang Jaringan Internet, Telekomunikasi, Artificial Intelligence, Internet Of Things (IoT), Augmented Reality (AR), Technopreneurship, Internet Marketing dan bidang pengolahan dan analisa data (komputer statistik).

Penulis adalah pendiri dari Universitas Sains dan Teknologi Komputer (Universitas STEKOM ) dan juga seorang dosen yang memiliki Jabatan Fungsional Akademik Lektor Kepala (Associate Professor) yang telah menghasilkan puluhan Buku Ajar ber ISBN, HAKI dari beberapa karya cipta dan Hak Paten pada produk IPTEK. Penulis juga terlibat dalam berbagai organisasi profesi dan industri yang terkait dengan dunia usaha dan industri, khususnya dalam pengembangan sumber daya manusia yang unggul untuk memenuhi kebutuhan dunia kerja secara nyata.



YAYASAN PRIMA AGUS TEKNIK

**PENERBIT :**

**YAYASAN PRIMA AGUS TEKNIK**

JL. Majapahit No. 605 Semarang  
Telp. (024) 6723456. Fax. 024-6710144  
Email : [penerbit\\_ypat@stekom.ac.id](mailto:penerbit_ypat@stekom.ac.id)

ISBN 978-623-5734-41-5 (PDF)



# **KETRAMPILAN PENALARAN DEDUKTIF**

**(Deductive reasoning skills)**

**Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM**



**YAYASAN PRIMA AGUS TEKNIK**

**PENERBIT :**

**YAYASAN PRIMA AGUS TEKNIK**

JL. Majapahit No. 605 Semarang

Telp. (024) 6723456. Fax. 024-6710144

Email : [penerbit\\_ypat@stekom.ac.id](mailto:penerbit_ypat@stekom.ac.id)

## **Ketrampilan Penalaran Deduktif (Deductive reasoning skills)**

### **Penulis :**

Dr. Ir. Agus Wibowo, M.Kom., M.Si., MM.

**ISBN : 9 786235 734415**

### **Editor :**

Dr. Joseph Teguh Santoso, S.Kom., M.Kom.

### **Penyunting :**

Dr. Mars Caroline Wibowo. S.T., M.Mm.Tech

### **Desain Sampul dan Tata Letak :**

Irdha Yudianto, S.Ds., M.Kom.

### **Penebit :**

Yayasan Prima Agus Teknik Bekerja sama dengan  
Universitas Sains & Teknologi Komputer (Universitas STEKOM)

### **Redaksi :**

Jl. Majapahit no 605 Semarang  
Telp. (024) 6723456  
Fax. 024-6710144  
Email : [penerbit\\_ypat@stekom.ac.id](mailto:penerbit_ypat@stekom.ac.id)

### **Distributor Tunggal :**

#### **Universitas STEKOM**

Jl. Majapahit no 605 Semarang  
Telp. (024) 6723456  
Fax. 024-6710144  
Email : [info@stekom.ac.id](mailto:info@stekom.ac.id)

Hak cipta dilindungi undang-undang  
Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara  
apapun tanpa ijin dari penulis

## KATA PENGANTAR

Puji syukur kepada Tuhan yang maha Esa, bahwa pada akhirnya buku yang berjudul "Ketrampilan Penalaran Deduktif (*Deductive reasoning skills*)" bisa diselesaikan dengan baik. Sejak zaman Plato (375 SM/2003), telah dikemukakan bahwa manusia dapat berpikir secara logis, dan matematis. Ini dikenal sebagai teori disiplin formal / *Theory of Formal Discipline* (TFD) dan dijelaskan oleh filsuf John Locke bahwa matematika harus diajarkan "kepada semua orang yang memiliki waktu dan kesempatan, agar mereka rasional, namun bukan sebagai ahli matematika."

Tujuan penulisan buku ini adalah untuk memaparkan hasil penelitian kebenaran TFD: apakah benar mempelajari matematika dapat mengembangkan penalaran berpikir, dan jika demikian, bagaimana caranya? Pada bab 1 mengulas psikologi teori penalaran dan penelitian sebelumnya yang terkait dengan TFD. Pada bab 2 membahas etika penelitian, metode eksperimen dan eksperimen semu, dan masalah reliabilitas dan validitas. Pada bab 3 diperkenalkan tugas-tugas yang paling umum yang digunakan untuk mengukur keterampilan penalaran, dan menyajikan argumen untuk menggunakan tugas inferensi kondisional Evans, Clibbens dan Rood (1995) sebagai ukuran utama dalam keterampilan penalaran buku ini.

Pada bab 5 buku ini disajikan dalam studi kuasi-eksperimental longitudinal yang akan menyelidiki perkembangan keahlian pemikiran kondisional dan silogistik pada mahasiswa matematika dibandingkan dengan mahasiswa sastra. Hal ini menunjukkan bahwa meskipun tidak ada perbedaan perilaku berpikir antar kelompok pada awal belajar, namun mahasiswa matematika mengubah interpretasi keadaan secara signifikan setelah semester satu selesai, sedangkan mahasiswa sastra tidak berubah. Namun, pemikiran mahasiswa tidak secara cepat menjadi normatif (yaitu, secara formal dan logis benar). Sebaliknya, mereka akan semakin mengadopsi dengan apa yang disebut interpretasi "cacat" dari klausa kondisional. Hal ini, pada awalnya dapat dilihat sebagai peningkatan interpretasi bikondisional yang lebih umum.

Apabila persyaratan "jika p maka q" disalahartikan, kasus non-p dianggap tidak relevan secara kondisional. Akibatnya, tidak dapat mendukung inferensi mode Tollens yang dianggap valid dalam model normatif logika kondisional (berbagai model kondisional juga dibahas dalam Bab 2). Dalam Bab 6, pola perubahan yang sama diamati untuk mahasiswa matematika, tetapi efeknya dapat diabaikan dalam penelitian karena dayanya yang berkurang. Pada Bab 7 disajikan studi eksperimental, di mana mahasiswa pascasarjana (level AS) diberi tugas inferensi kondisional dengan pernyataan kondisional yang diutarakan sebagai 'jika p maka q' atau setara secara logis 'p hanya jika q'. Kelompok '*hanya jika*' tampil secara signifikan kurang sejalan dengan cacat kondisional daripada kelompok '*jika kemudian*', menunjukkan bahwa perubahan interpretasi yang ditemukan dalam studi tingkat sarjana adalah khusus untuk ungkapan linguistik yang biasa ditemui dalam matematika '*jika p maka Q*'. Hal ini bertentangan dengan asumsi TFD bahwa belajar matematika akan membawa perubahan luas dalam perilaku pemikiran.

Pada bab 8 menggunakan versi waktu terbatas dari masalah penalaran kondisional untuk menunjukkan bahwa perbedaan dalam penerimaan disabilitas antara ahli matematika dan non-matematika sebagian disebabkan karena pemrosesan kognitif otomatis, tetapi tidak

seluruhnya. Akhirnya, pada bab 9 meneliti hubungan antara perilaku penalaran kondisional dan fungsi eksekutif - keterampilan yang memungkinkan kita untuk mengontrol perhatian dan upaya kognitif kita dan menyarankan bahwa keterampilan yang menghambat dan menggeser mungkin lebih memainkan peran dalam mengadopsi interpretasi kondisional yang cacat atas penafsiran bikondisional.

Pada bagian akhir buku ini diakhiri dengan bab 9, dengan proposal bahwa belajar matematika membuat mahasiswa pada pernyataan implisit dalam bentuk 'jika p maka q', di mana mereka diharapkan untuk mengasumsikan jika p benar dan menyimpulkan sesuatu tentang q . Hal ini nampaknya mendorong interpretasi yang cacat dari kondisional yang terbatas pada penalaran kondisional abstrak dari bentuk 'jika p maka q'. Oleh karena itu, hubungan antara mempelajari matematika dan perubahan keterampilan penalaran mungkin jauh lebih terbatas daripada yang diperkirakan sebelumnya. Akhir kata semoga buku ini berguna bagi para pembaca.

Semarang, Februari 2022

Dr. Ir Agus Wibowo, M.Kom, M.Si, MM

## DAFTAR ISI

<b>Halaman Judul</b> .....	<b>i</b>
<b>Kata Pengantar</b> .....	<b>iii</b>
<b>Daftar Isi</b> .....	<b>v</b>
<b>BAB 1 DASAR TEORI PENALARAN</b> .....	<b>1</b>
1.1 Pendahuluan .....	1
1.2 Mendefinisikan Penalaran .....	1
1.3 Rasionalitas – Penalaran apa yang “Lebih Baik”? .....	4
1.4 Sejarah Singkat Teori Disiplin Formal .....	10
1.5 Penelitian keterampilan bernalar dan hubungannya dengan matematika .....	12
1.6 Psikologi Penalaran .....	20
1.7 Status Teori Disiplin Formal Saat Ini .....	34
<b>BAB 2 DASAR METODE PENELITIAN</b> .....	<b>37</b>
2.1 Pendahuluan .....	37
2.2 Gambaran Umum Studi Longitudinal .....	37
2.3 Etika Penelitian .....	37
2.4 Ilmu Pengetahuan dan Masyarakat .....	37
2.5 Kesejahteraan Partisipan .....	38
2.6 Metode Eksperimen .....	38
2.7 Metode Eksperimen Semu .....	41
2.8 Reabilitas dan Validitas .....	45
<b>BAB 3 MENGUKUR PENALARAN</b> .....	<b>49</b>
3.1 Pendahuluan .....	49
3.2 Penghakiman dan Pengambilan Keputusan .....	49
3.3 Tugas Heuristik dan Bias .....	49
3.4 Pemikiran Deduktif .....	53
3.5 Tugas DisjungSIONAL .....	54
3.6 Tugas Kondisional .....	57
3.7 Tugas Silogisme .....	61
3.8 Ringkasan .....	63
<b>BAB 4 MENGEMBANGKAN KETRAMPILAN BERNALAR PADA MAHASISWA</b> .....	<b>65</b>
4.1 Pengujian Teori Disiplin Formal .....	65
4.2 Ringkasan .....	67
4.3 Studi Pilot .....	68
4.4 Ringkasan Studi Pilot .....	75
4.5 Studi Utama .....	75
4.6 Metode .....	75
4.7 Hasil Pengukuran .....	76
4.8 Diskusi .....	89
4.9 Pengembangan Keterampilan Penalaran .....	89

4.10 Kompetensi dan bias dalam Tugas Inferensi Kondisional .....	90
4.11 Keterbatasan .....	91
4.12 Revisi status Teori Disiplin Formal .....	92
4.13 Ringkasan dan temuan baru .....	92
<b>BAB 5 MENGENAL KEMAMPUAN BERPIKIR PASCASARJANA .....</b>	<b>94</b>
5.1 Menguji Teori Disiplin Formal .....	94
5.2 Ringkasan .....	95
5.3 Metode .....	96
5.4 Pengukuran .....	96
5.5 Prosedur .....	98
5.6 Hasil .....	98
5.7 Analisis Awal Kovariat .....	99
5.8 Pengembangan Keahlian Pemikiran .....	100
5.9 Diskusi .....	106
5.10 Ringkasan dan temuan Baru .....	107
<b>BAB 6 FAKTOR BAHASA MATEMATIKA: 'JIKA KEMUDIAN' DAN 'HANYA JIKA' .....</b>	<b>108</b>
6.1 Pendahuluan .....	108
6.2 Metode .....	110
6.3 Partisipan .....	110
6.4 Tarif Pengesahan .....	111
6.5 Interpretasi Kondisi .....	112
6.6 Diskusi .....	112
6.7 Ringkasan dan temuan Baru .....	115
<b>BAB 7 PERAN HEURISTIK DALAM PENALARAN MAHASISWA .....</b>	<b>116</b>
7.1 Pendahuluan .....	116
7.2 Metode .....	117
7.3 Hasil .....	118
7.4 Analisis Utama .....	119
7.5 Diskusi .....	124
7.6 Ringkasan dan temuan Baru .....	126
<b>BAB 8 PERAN FUNGSI EKSEKUTIF DALAM BERPIKIR KONDISIONAL .....</b>	<b>127</b>
8.1 Pendahuluan .....	127
8.2 Metode .....	130
8.3 Pengukuran .....	130
8.4 Prosedur .....	132
8.5 Hasil .....	132
8.6 Performa Tugas .....	132
8.7 Hubungan di antara Fungsi Eksekutif .....	133
8.8 Fungsi Eksekutif dan Penalaran Kondisional. ....	133
8.9 Diskusi .....	134
8.10 Ringkasan dan temuan Baru .....	136
<b>BAB 9 SIMPULAN .....</b>	<b>137</b>
9.1 Pendahuluan .....	137



9.2 Gambaran Umum Interpretasi dan Penemuan .....	137
9.3 Penelitian Lanjutan .....	140
9.4 Meninjau Kembali Teori Displin Formal .....	141
<b>DAFTAR LAMPIRAN .....</b>	<b>142</b>
<b>DAFTAR PUSTAKA .....</b>	<b>156</b>

## BAB 1 DASAR TEORI PENALARAN

### 1.1 PENDAHULUAN

Buku ini bertujuan untuk membahas klaim yang dikenal sebagai *Theory of Formal Discipline* (TFD): bahwa belajar matematika meningkatkan keterampilan penalaran umum ke tingkat yang lebih besar daripada disiplin ilmu lainnya. Pembahasan tentang masalah ini terbagi menjadi dua bagian:

1. Apakah mempelajari matematika dapat meningkatkan keterampilan bernalar secara lebih baik daripada mengikuti mata kuliah lain.
2. Mekanisme kognitif apa yang terkait dengan keterampilan penalaran 'yang lebih baik' dan apakah ini bertanggung jawab atas peningkatan yang ditemukan di bagian 1?

Tinjauan literatur yang disajikan di bawah ini akan mendefinisikan penalaran, membahas beberapa perspektif tentang rasionalitas, meninjau sejarah TFD dan bukti yang berkaitan dengan klaimnya, membahas beberapa teori penalaran dan bagaimana hal itu dapat ditingkatkan, dan bagian terakhir, ini akan merangkum status terkini dari TFD dan mengapa ini penting untuk diselidiki.

### 1.2 MENDEFINISIKAN PENALARAN

Sebelum membahas literatur tentang bagaimana kita bernalar dan hubungannya dengan matematika, penting untuk memperjelas apa itu penalaran. Penalaran dapat dilihat sebagai proses kognitif untuk menyimpulkan informasi baru dari informasi yang diberikan, dan secara umum ada dua bentuk: deduktif dan induktif. Dalam penalaran deduktif, suatu kesimpulan harus benar jika premis-premisnya benar, misalnya, 'jika  $p$  maka  $q$ '. Dengan penalaran induktif, sebuah kesimpulan mungkin benar ketika premisnya benar, misalnya, 'semua  $ps$  yang terlihat sejauh ini adalah  $qs$ ; karena itu; semua  $ps$  adalah  $qs$ '. Sehingga, penalaran deduktif lebih ketat daripada penalaran induktif, dan ini merupakan bentuk yang digunakan dalam pembuktian matematis. Faktanya, Polya (1954, p. V) menyatakan bahwa "[a] bukti matematis adalah penalaran [deduktif]".

<p>Jika <math>p</math> maka <math>q</math>  <math>p</math>            Kesimpulan : <math>q</math>            a) Modus Ponens (MP)</p>
<p>Jika <math>p</math> maka <math>q</math>  <math>q</math>            Kesimpulan : <math>p</math>            c) Affirmation of the consequent/Penegasan akibat (AC)</p>
<p>Jika <math>p</math> maka <math>q</math>            bukan <math>p</math>            Kesimpulan : bukan <math>q</math>            b) Penolakan pendahulunya/Peolakan Sebelumnya (DA)</p>
<p>Jika <math>p</math> maka <math>q</math>            bukan <math>q</math>            Kesimpulan : bukan <math>p</math>            d) Affirmation of the consequent/Penegasan akibat (AC)</p>

**Gambar 1.1:** Struktur dasar dari empat kesimpulan: Modus Ponens, Denial of the Anteseden/penplakan sebelumnya, Armation of the Consequent/penegasan akibat, dan Modus Tollens.

Penalaran deduktif dapat dibagi menjadi beberapa bentuk, seperti penalaran kondisional, disjungtif dan silogistik, dan ini dibahas secara lebih rinci dalam Bab 4. Namun, penalaran kondisional akan menjadi pusat penelitian sehingga perlu diuraikan di sini.

Penalaran kondisional adalah proses menarik kesimpulan dari pernyataan kondisional seperti 'jika p maka q'. Ini dianggap sebagai pusat logika (Anderson & Belnap, 1975; Braine, 1978; Inglis & Simpson, 2008) dan matematika (Houston, 2009). Ada empat kesimpulan yang biasanya diambil dari pernyataan kondisional: modus ponens (MP), modus tollens (MT), penolakan antecedent/ Denial of the Antecedent (DA) dan penegasan konsekuensi/ Affirmation of the Consequent (AC). Struktur masing-masing inferensi ini ditunjukkan pada Gambar 1.1.

Ada juga empat bentuk pernyataan kondisional yang dibuat oleh ada dan tidak adanya negasi: 'jika p maka q', 'jika p maka bukan q', 'jika tidak p maka q', dan 'jika tidak p maka tidak q'. Tabel 1.1 menunjukkan premis dan kesimpulan dari setiap kombinasi pernyataan kondisional dan tipe inferensi. Untuk menguraikan contoh dari tabel, inferensi AC dari pernyataan 'jika tidak p maka q' akan berbunyi:

*Aturan: Jika tidak p maka q*

*Premis: q*

*Kesimpulan: Oleh karena itu, bukan p*

**Tabel 1.1:** Empat jenis inferensi dan pernyataan kondisional dengan dan tanpa premis negasi (Pr) dan kesimpulan (Con). Simbol harus dibaca 'tidak'.

<i>Kondisional</i>	<b>MP</b>		<b>DA</b>		<b>AC</b>		<b>MT</b>	
	<i>Pr</i>	<i>Con</i>	<i>Pr</i>	<i>Con</i>	<i>Pr</i>	<i>Con</i>	<i>Pr</i>	<i>Con</i>
<i>jika p maka q</i>	<i>p</i>	<i>q</i>	<i>¬p</i>	<i>¬q</i>	<i>q</i>	<i>P</i>	<i>¬q</i>	<i>¬p</i>
<i>jika p maka ¬q</i>	<i>p</i>	<i>¬q</i>	<i>¬p</i>	<i>q</i>	<i>¬q</i>	<i>P</i>	<i>q</i>	<i>¬p</i>
<i>jika ¬p maka q</i>	<i>¬p</i>	<i>q</i>	<i>p</i>	<i>¬q</i>	<i>Q</i>	<i>¬p</i>	<i>¬q</i>	<i>P</i>
<i>jika ¬p maka ¬q</i>	<i>¬p</i>	<i>¬q</i>	<i>p</i>	<i>q</i>	<i>¬q</i>	<i>¬p</i>	<i>q</i>	<i>p</i>

Validitas masing-masing dari empat kesimpulan tergantung pada cara seseorang menafsirkan pernyataan kondisional. Ada empat interpretasi umum dari sebuah kondisi: pembacaan kondisional material, bikondisional, kondisional cacat dan konjungtif (Evans, Handley, Neilens & Over, 2007). Masing-masing ditunjukkan dalam bentuk tabel kebenaran pada Tabel 1.2.

Di bawah pembacaan kondisional material, yang dianggap benar oleh ahli logika, pernyataan 'jika p, maka q' berarti bahwa p cukup untuk q, dan q diperlukan untuk p. Dengan kata lain, conditional berarti 'q atau bukan-p'. Kondisional benar dalam semua kasus kecuali jika p benar dan q salah. Ada kemungkinan p salah dan q benar, karena p belum tentu satu-satunya determinan q.

Di bawah pembacaan bikondisional, kondisi diperlakukan sebagai 'p jika dan hanya jika q'. Dengan kata lain, p menyiratkan q dan q menyiratkan p, sehingga keduanya diperlukan dan cukup untuk satu sama lain. Dengan demikian, pernyataan 'jika p, maka q' tidak benar ketika p salah dan q benar serta ketika p benar dan q salah – p dan q keduanya harus benar atau keduanya salah.

Di bawah pembacaan kondisional yang cacat, p dianggap perlu dan cukup untuk q, jadi hanya kasus p q yang dianggap membuat aturan itu benar. Namun, alih-alih kasus bukan-p yang membuat aturan menjadi salah, kasus-kasus tersebut dianggap tidak relevan. Aturan dianggap benar jika p dan q keduanya benar, salah jika p benar dan q salah, dan tidak relevan jika p tidak benar. Hal ini mengarah pada pola kesimpulan yang serupa yang dideduksi seperti

dalam kasus kondisional material, kecuali untuk MT. Contoh inferensi MT adalah sebagai berikut:

*Jika p maka q*  
*bukan-q*  
*Oleh karena itu tidak-p*

**Tabel 1.2:** Tabel kebenaran untuk empat interpretasi 'jika p maka q' di mana t = benar, f = salah, dan i = tidak relevan.

<i>p</i>	<i>q</i>	<i>Material</i>	<i>Kondisional</i>	<i>Defektif</i>	<i>Konjungtif</i>
<i>t</i>	<i>t</i>	<i>t</i>	<i>t</i>	<i>t</i>	<i>t</i>
<i>t</i>	<i>f</i>	<i>f</i>	<i>f</i>	<i>f</i>	<i>f</i>
<i>f</i>	<i>t</i>	<i>t</i>	<i>f</i>	<i>i</i>	<i>f</i>
<i>f</i>	<i>f</i>	<i>t</i>	<i>t</i>	<i>i</i>	<i>f</i>

Inferensi MT dianggap sah di bawah interpretasi material tetapi tidak valid di bawah interpretasi yang rusak. Ini karena premis minor tidak mendukung p dan karena itu dianggap tidak relevan dengan makna kondisional sehingga tidak ada kesimpulan yang dapat ditarik. Namun, jika penalaran mampu membangun bukti kontradiksi maka mereka mungkin dapat menyimpulkan notp dari premis notq dan menerima inferensi MT. Misalnya "anggap kontradiksi p, simpulkan q, tapi saya tidak tahuq, jadi asumsi p pasti salah". Proses ini membutuhkan memori kerja tingkat tinggi dan mungkin hanya tersedia bagi para pemikir yang paling cakap. Untuk alasan ini, MT umumnya dianggap tidak valid jika disalahartikan. Akhirnya, dalam bacaan asosiatif, pernyataan kondisional ditafsirkan sebagai "p dan q", sehingga pernyataan kondisional benar hanya jika p dan q keduanya benar dan salah jika tidak.

Perhatikan contoh 'jika Amerika ada di Eropa, maka Indonesia ada di Asia'. Kita tahu bahwa anteseden salah dan konsekuen benar, dan kita dapat melihat dari Tabel 1.2 bahwa kasus konsekuen benar anteseden salah dianggap tidak relevan dengan aturan di bawah interpretasi yang cacat, tetapi dianggap membuat aturan benar di bawah materi kondisional dan salah di bawah interpretasi bikondisional dan konjungtif.

Tabel 1.3 menunjukkan apakah masing-masing dari empat kesimpulan (MP, MT, DA, AC) dianggap valid di bawah masing-masing dari empat interpretasi kondisional. Evans dkk. (2007) menemukan bahwa MP hampir diterima secara universal (97,5%), diikuti oleh AC (74%), MT (50%) dan terakhir DA (38,5%). Ini adalah kondisi material yang dianggap benar oleh ahli logika, tetapi peserta Evans et al. jelas tidak memegang interpretasi ini. Kondisional dalam bahasa sehari-hari cenderung mengasumsikan interpretasi bikondisional atau cacat (Cummins, 1995; Markovits, 1985; Venet & Markovits, 2001; Zepp, 1987; Verschueren, Schaeken & d'Ydewalle, 2005) dan seterusnya secara keseluruhan, orang tidak berpengalaman dengan materi kondisional, seperti yang tercermin dalam tingkat dukungan dari peserta Evans et al (2007). Kemungkinan yang menarik dalam konteks penelitian ini adalah studi matematika meningkatkan keakraban dan kompetensi seseorang dengan materi kondisional, yang mengarah ke kinerja yang lebih normatif pada Tugas Inferensi Kondisional.

**Tabel 1.3:** Validitas empat inferensi di bawah setiap interpretasi kondisional.

Interfensi	Material	Kondisional	Defektif	Konjungtif
MP	Valid	Valid	Valid	Valid
DA	Invalid	Valid	Invalid	Invalid
AC	Invalid	Valid	Invalid	Invalid
MT	Valid	Valid	Invalid	Invalid

Kemampuan inferensi kondisional akan menjadi fokus penelitian ini untuk alasan yang dibahas dalam bab 3, tetapi tinjauan pustaka akan mempertimbangkan penelitian yang relevan ke dalam segala bentuk penalaran karena, seperti yang ditunjukkan di bawah ini, klaim TFD tidak jelas tentang keterampilan mana yang ditingkatkan oleh mempelajari matematika. Setiap jenis penalaran lebih lanjut yang dibahas akan didefinisikan saat diperkenalkan.

### 1.3 RASIONALITAS – PENALARAN APA YANG “LEBIH BAIK”?

Buku ini berkaitan dengan apakah penalaran dapat ditingkatkan, sehingga masalah yang jelas adalah apa yang harus dianggap 'lebih baik' dan 'buruk' dari penalaran. Terbukti secara luas bahwa perilaku penalaran orang bias (yaitu menyimpang dari model penalaran normatif, Kahneman & Tversky, 1972; Stanovich, 2009b), tetapi masalahnya adalah apakah ini menunjukkan irasionalitas atau apakah itu dapat dijelaskan dan dimaafkan dalam beberapa hal lain.

Beberapa orang berpendapat bahwa bias yang cenderung kita alami bukanlah bukti penalaran yang buruk sama sekali, sementara tentu saja yang lain tidak setuju dan berpendapat bahwa kita bisa melakukan yang lebih baik. Sebelum mengelaborasi perspektif-perspektif ini, beberapa bias penalaran yang umum didiskusikan untuk memperjelas masalah yang ditentang oleh pandangan-pandangan yang berbeda ini.

#### 1.3.1 Heuristik dan Bias

Ada banyak literatur tentang heuristik dan bias dalam perilaku penalaran, dan bagian ini hanya akan mengulas beberapa masalah umum. Istilah heuristik mengacu pada 'jalan pintas' kognitif, yang digunakan sebagai pengganti proses bernalar yang lebih lama, dan yang mungkin disadari atau tidak disadari (Gigerenzer & Gaissmaier, 2011). Heuristik menghemat banyak upaya mental dan biasanya sangat efektif, tetapi dalam beberapa kasus mereka dapat menyebabkan bias.

Bias adalah penyimpangan sistematis dari model rasionalitas normatif: kesalahan yang cenderung ditunjukkan berulang kali di dalam dan di antara para pemikir (Kahneman & Tversky, 1972). Bias dianggap kesalahan berdasarkan asumsi bahwa model normatif memaksimalkan hasil individu, dan jadi jika seseorang tidak mengikuti model normatif dari situasi penalaran yang diberikan, itu merugikan mereka sendiri (Stanovich, 1999). Bab 4 memberikan pengantar tugas yang paling umum digunakan untuk mengukur perilaku penalaran, dan dalam melakukannya juga memberikan gambaran tentang banyak bias yang umum diamati. Di sini, hanya beberapa bias yang diperkenalkan untuk menunjukkan bahwa manusia tidak selalu berperilaku sesuai dengan model normatif.

Dimulai dengan Tugas Inferensi Kondisional, dua bias yang umum diamati adalah bias kesimpulan negatif/ *negative conclusion bias* (NCB) dan bias premis afirmatif/ *affirmative premise bias* (APB). NCB mengacu pada kecenderungan peserta untuk menerima lebih banyak kesimpulan yang menghasilkan kesimpulan negatif daripada kesimpulan afirmatif. Misalnya,

inferensi DA 'jika p maka q, bukan p, maka bukan q' akan diterima lebih sering daripada inferensi DA 'jika p maka bukan q, bukan p, oleh karena itu q' meskipun secara logika ekuivalen (Evans et al. , 1995; Evans & Handley, 1999). Efek ini diamati pada kedua kesimpulan penolakan (MT dan DA) tetapi hanya kadang-kadang dan lemah pada AC, dan tidak pernah pada MP (Schroyens, Schaeken & d'Ydewalle, 2001). Ada dua penjelasan populer untuk NCB.

Satu saran adalah orang berasumsi bahwa bukan-p lebih umum daripada p di dunia nyata (misalnya ada lebih banyak hal tidak biru daripada ada hal biru), dan karena itu lebih bersedia untuk menyimpulkan bukan-p daripada p (Pollard & Evans, 1980; Oaksford, Chater & Larkin, 2000). Akun alternatif adalah NCB berasal dari masalah dengan pemrosesan negasi ganda. NCB paling sering diamati pada inferensi MT dan DA, di mana kesimpulan afirmatif dihasilkan dari negasi ganda, misalnya 'jika bukan A maka 3; bukan 3; oleh karena itu bukan (bukan A)'. Penurunan tingkat dukungan dalam masalah ini mungkin karena alasan gagal untuk mengubah 'bukan (bukan A)' menjadi 'A'.

APB mengacu pada pengamatan bahwa peserta lebih cenderung menerima kesimpulan dengan premis afirmatif daripada premis penolakan, terutama ketika penolakan itu tersirat (seperti '7' bukannya 'bukan 3', Evans et al., 1995). Misalnya, inferensi 'jika A bukan 5, maka 5 bukan A' akan diterima lebih sering daripada inferensi 'jika A bukan 5, maka A adalah 8', meskipun keduanya merupakan inferensi AC yang tidak valid. Ini telah dijelaskan sebagai bias yang cocok (lihat lebih lanjut di bawah), di mana premis '5' lebih jelas terkait dengan kondisional daripada premis '8' (Evans & Handley, 1999).

Bias lain yang sangat relevan dengan penelitian ini adalah bias keyakinan. Ini adalah kecenderungan peserta untuk bernalar menurut keyakinan mereka sebelumnya daripada informasi yang ada (Evans, Barston & Pollard, 1983; Sa, West & Stanovich, 1999). Misalnya, ketika berhadapan dengan silogisme "segala sesuatu dengan empat kaki berbahaya; Anjing tidak berbahaya; oleh karena itu, anjing tidak memiliki empat kaki", seseorang yang bias oleh keyakinan mereka sebelumnya bahwa anjing memiliki empat kaki akan menjawab bahwa silogisme tidak valid, sedangkan orang yang mendasarkan penilaiannya pada informasi yang ada akan menjawab bahwa itu valid. Hanya pada item di mana validitas dan kepercayaan berada dalam konflik (yaitu masalah yang valid dan tidak dapat dipercaya atau tidak valid dan dapat dipercaya) di mana bias kepercayaan dapat ditampilkan.

Bagian ini hanya menjelaskan sedikit pilihan bias yang telah ditunjukkan secara empiris dalam literatur penalaran, tetapi cukup untuk menunjukkan bahwa orang tidak selalu bernalar menurut standar normatif. Bagian berikut membahas perbedaan ini dalam hal berbagai pandangan tentang rasionalitas dan apakah kita dapat mengharapkan orang untuk mampu 'berbuat lebih baik' dengan penalaran menurut model normatif.

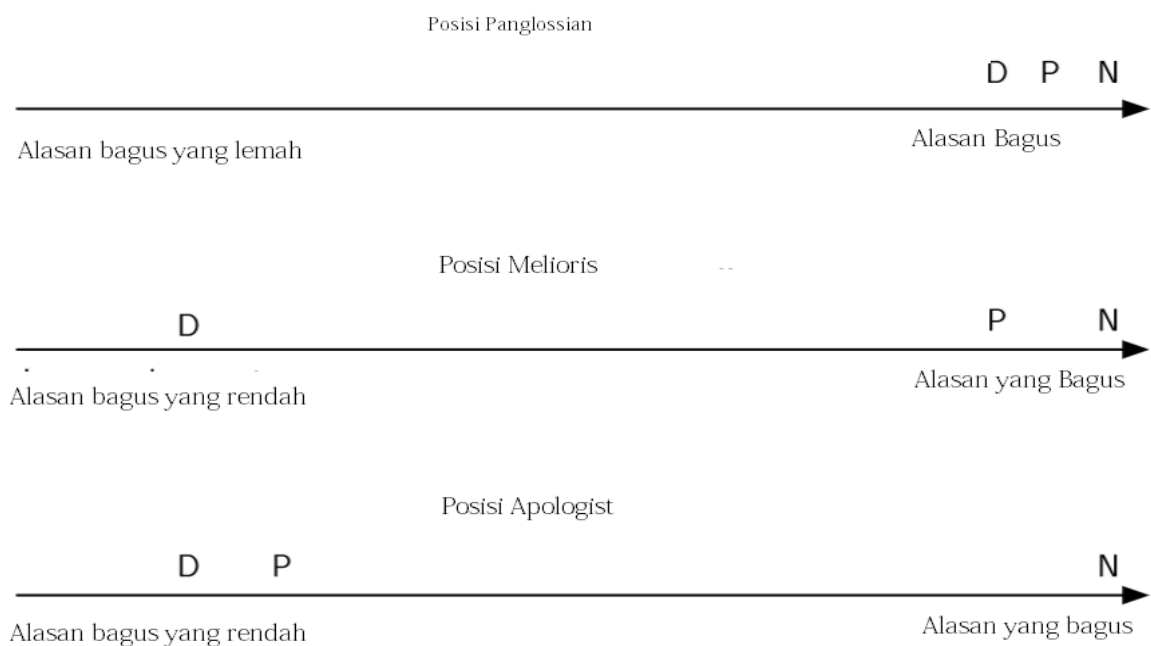
### **1.3.2 Apakah Manusia Tidak Rasional?**

Seperti yang ditunjukkan di atas, manusia sering menunjukkan bias ketika bernalar. Masalah yang dihadapi adalah apakah bias ini menunjukkan irasionalitas dan apakah masuk akal untuk mengharapkan orang berbuat lebih baik, atau setidaknya mampu berbuat lebih baik. Hal ini adalah pertanyaan penting dalam sebuah penelitian, karena jika kita tidak dapat mengharapkan orang untuk berbuat lebih baik, maka pertanyaan apakah matematika secara khusus membuat orang lebih baik dalam penalaran akan menjadi mubazir.

Ada tiga posisi rasionalitas manusia yang berbeda dalam perspektifnya tentang hubungan antara model penalaran manusia deskriptif, preskriptif dan normatif (Stanovich, 1999). Model deskriptif menggambarkan dan berteori tentang pola penalaran manusia. Model

normatif menetapkan standar ideal untuk penalaran, yang jika dicapai akan memaksimalkan hasil bagi individu. Namun, seperti yang dikatakan Harman (1995) dan Stich (1990), manusia memiliki kecerdasan yang terbatas dan waktu yang terbatas, dan tidak masuk akal untuk mengharapkan kita bertindak dengan cara yang rasional secara normatif mengingat pembatasan ini. Argumen inilah yang mengarah pada gagasan model preskriptif. Model preskriptif menentukan yang terbaik yang bisa kita harapkan untuk dicapai saat bernalar, mengingat keterbatasan kognitif dan seringkali situasional yang harus dikerjakan oleh si pemikir. Ini kemudian menggantikan model normatif sebagai standar yang harus kita harapkan untuk dicapai dan dengan mana kita harus membandingkan model deskriptif.

Seperti disebutkan di atas, ada tiga posisi pada hubungan antara deskriptif, preskriptif dan model normatif perilaku penalaran, dan ini ditunjukkan pada Gambar 1.2. *'Panglossians'* tidak melihat bahwa ada kesenjangan substansial antara ketiga model (Stanovich, 1999). Mereka berpendapat bahwa manusia bernalar sebaik mungkin dan sebaik yang seharusnya, dan oleh karena itu irasionalitas manusia tidak menjadi masalah. Bias yang dijelaskan di atas dijelaskan sebagai kesalahan kinerja acak, penerapan norma yang salah di pihak eksperimen, atau kesalahpahaman masalah oleh peserta, karena eksperimen tidak jelas (Stanovich & West, 2000). Penjelasan-penjelasan tersebut diuraikan di bawah ini. Dalam situasi apa pun bias dianggap sebagai demonstrasi irasionalitas pada peserta.



**Gambar 1.2:** Tiga posisi rasionalitas manusia. N = model normatif, P = model preskriptif, dan D = model deskriptif.

*'Meliorists'*, di sisi lain, berpendapat bahwa cara kita sebenarnya bernalar jauh dari standar yang dengannya kita dapat dan harus bernalar. Mereka melihat model preskriptif sebagai dekat dengan model normatif, sedangkan model deskriptif agak jauh dari dua lainnya. Kesenjangan antara model deskriptif dan preskriptif karena itu dapat dianggap irasionalitas, karena kita gagal untuk melakukan sebaik yang kita bisa diharapkan untuk melakukannya (Stanovich, 1999).

Posisi terakhir adalah '*Apologist*'. Posisi Apologis setuju dengan Meliorists dalam model deskriptif jauh dari model normatif, tetapi berbeda dalam menempatkan model preskriptif lebih dekat dengan deskriptif. Maka, ia berpendapat bahwa sementara kita tidak menalar dengan sempurna, kita melakukan yang terbaik yang bisa kita harapkan mengingat keterbatasan kognitif kita (Stanovich, 1999). Seperti Panglossians, Apologis percaya bahwa kita tidak bisa menuduh manusia irasionalitas, karena kita hanya serasional mungkin. Perbedaan antara keduanya adalah Apologis mengakui bahwa ini tidak sesuai dengan standar ideal/normatif, sedangkan Panglossian berpendapat demikian.

Sepanjang bukunya '*Who is Rational?*', Stanovich (1999) menggunakan data tentang perbedaan individu dalam perilaku penalaran untuk menyelidiki kemungkinan alasan untuk kesenjangan yang jelas antara model deskriptif dan normatif. Alasan yang mungkin, seperti disinggung di atas, adalah: kesalahan kinerja acak, penerapan norma yang salah, konstruksi tugas alternatif, keterbatasan komputasi, dan irasionalitas sistematis.

Argumen Panglossian yang digunakan untuk mempertahankan rasionalitas manusia adalah kesenjangan yang tampak antara model normatif dan deskriptif disebabkan oleh kesalahan kinerja acak oleh para peserta – mungkin mereka sejenak terganggu, misalnya (Stanovich & West, 2000). Masalah yang diidentifikasi Stanovich (1999) dengan penjelasan ini adalah kesalahan bersifat sistematis, bukan acak. Orang-orang secara konsisten membuat kesalahan yang sama, dan tingkat kesalahan mereka pada satu tugas memprediksi tingkat kesalahan mereka pada tugas lain. Kinerja juga terkait dengan variabel kognitif dan kepribadian, seperti kecerdasan (atau kapasitas kognitif), yang bertentangan dengan pandangan kesalahan acak. Sementara semua ukuran yang digunakan dalam penelitian psikologis pasti akan tunduk pada beberapa kesalahan acak, atau 'kebisingan', ini bukan penjelasan yang cukup untuk menjelaskan kesenjangan substansial antara model penalaran deskriptif dan normatif.

Argumen norma yang salah menempatkan kesalahan atas kesenjangan normatif/deskriptif dengan pelaku eksperimen (Stanovich, 1999; Stanovich & West, 2000). Dalam literatur heuristik dan bias, kinerja biasanya dibandingkan dengan norma statistik atau logika, dan dapat dikatakan bahwa ini tidak tepat. Beberapa akan menyarankan bahwa alih-alih menggunakan norma-norma yang berasal dari statistik atau logika, tanggapan yang diberikan paling sering oleh peserta sebenarnya harus dianggap norma, karena orang pada dasarnya rasional (Stanovich, 1999). Stanovich, bagaimanapun, berpendapat bahwa jika kita ingin menggunakan kinerja peserta untuk menentukan norma, kita harus melihat tanggapan dari yang paling cerdas.

Kecerdasan dapat dianggap sebagai kemampuan yang konsisten untuk efektivitas dalam lingkungan dan situasi yang berbeda. Hal ini menunjukkan bahwa perilaku individu yang paling cerdas mencerminkan perilaku yang 'lebih baik' atau lebih efektif daripada individu yang kurang cerdas. Ketika perilaku individu yang paling cerdas sejalan dengan model normatif yang ditetapkan oleh para ahli, mungkin ini bukti bahwa model normatif memang sesuai. Jika, di sisi lain, individu yang paling cerdas memberikan respons non-normatif, maka mungkin kita harus mempertimbangkan untuk merevisi model normatif seperti yang disarankan beberapa orang.

Stanovich (1999) berpendapat bahwa skenario pertama biasanya terjadi. Dia menunjukkan bahwa di berbagai tugas, peserta konsisten apakah mereka memberikan tanggapan normatif atau non-normatif dan bahwa ada korelasi yang signifikan antara kinerja



normatif dan kecerdasan umum atau disposisi bernalar. Dia berpendapat bahwa jika para ahli dan peserta yang lebih mampu setuju bahwa respons normatif adalah respons yang benar, maka kemungkinan besar itu adalah respons yang benar, dan ini terjadi di sebagian besar tugas yang dia teliti. Namun, ada sebagian kecil tugas di mana hal ini tidak terjadi dan dalam kasus seperti itu, mungkin ada ruang untuk argumen norma yang salah untuk menjelaskan perbedaan antara model normatif dan deskriptif.

Penjelasan lain yang diusulkan untuk melestarikan rasionalitas manusia adalah peserta salah mengartikan tugas yang diberikan kepada mereka. Eksperimen mungkin menerapkan norma yang benar untuk tugas yang mereka inginkan, dan partisipan mungkin memproses tanpa kesalahan, namun masih ada masalah karena partisipan merespons dengan benar interpretasi tugas yang berbeda. Cosmides dan Tooby (1992) berpendapat bahwa manusia memiliki algoritme bawaan untuk menyelesaikan tugas-tugas penalaran yang telah dihadapi sepanjang sejarah evolusi kita, dan bahwa ketika masalah dibingkai untuk memperoleh algoritme ini, kita sebenarnya bernalar dengan cukup baik.

Namun, Stanovich (1999) menggunakan argumen yang sama untuk masalah norma yang salah untuk menunjukkan bahwa konstruksi tugas normatif, dalam banyak kasus, harus dianggap benar. Mereka yang memiliki kemampuan kognitif umum yang lebih tinggi cenderung menafsirkan tugas dengan cara yang dimaksudkan oleh pelaku eksperimen. Sekali lagi, ini tidak berlaku untuk setiap tugas sehingga ada beberapa ruang untuk argumen bahwa tugas disajikan dengan cara yang ambigu, tetapi ini adalah kasus untuk sebagian besar tugas.

Orang bisa berargumen bahwa tidak benar menggunakan perilaku individu berkemampuan tinggi untuk membenarkan penerapan norma tradisional dan interpretasi tugas. Argumen tampaknya agak melingkar – individu berkemampuan tinggi dianggap berkemampuan tinggi karena mereka cenderung memberikan tanggapan normatif pada tes kecerdasan, dan oleh karena itu kami berasumsi bahwa karena mereka juga memberikan tanggapan normatif untuk tugas lain, norma harus benar. Namun, ada alasan yang sangat kuat untuk mempertimbangkan tanggapan individu dengan kemampuan kognitif tinggi sebagai 'lebih baik' – orang-orang ini mendapatkan gelar yang lebih baik (Farsides & Woodfield, 2003), memiliki kinerja dan kesuksesan pekerjaan yang lebih tinggi (Deary, 2001 Judge, Higgins, Thoresen & Barrick, 1999), pendapatan yang lebih tinggi (Ashenfelter & Rouse, 1999), dan, yang mengejutkan, mereka hidup lebih lama (Deary, 2008). Seperti yang dicatat oleh Frederick (2005), masalah ini dapat diklarifikasi dengan memberikan tanggapan dan skor kecerdasan peserta sebelumnya kepada peserta baru. Jika peserta baru dipengaruhi oleh tanggapan individu dengan kecerdasan tinggi, itu akan menunjukkan beberapa tingkat konsensus dalam gagasan bahwa individu cerdas cenderung membuat pilihan yang lebih baik dan bahwa kita ingin melakukan apa yang mereka lakukan.

Gagasan keterbatasan kognitif menjadi penyebab kesenjangan penalaran memiliki beberapa dukungan, meskipun itu bukan penjelasan yang lengkap. Kecerdasan terbatas dan terkait dengan kinerja penalaran (Evans et al., 2007; S'á et al., 1999; Stanovich & West, 2008). Namun, dalam analisis Stanovich (1999), setelah kecerdasan dikendalikan, ada varians substansial yang tersisa dalam kinerja penalaran, dan ini terkait lintas tugas dan dapat diprediksi dari disposisi bernalar individu (West, Toplak & Stanovich, 2008).

Disposisi bernalar mengacu pada gaya kognitif individu - ciri-ciri seperti sejauh mana mereka bersedia untuk berusaha memecahkan tugas, atau untuk mengubah keyakinan mereka berdasarkan bukti baru atau kurangnya koherensi. Bukti menunjukkan bahwa

sementara ada beberapa kesalahan untuk kesenjangan deskriptif/normatif yang ditempatkan pada keterbatasan kognitif, ada juga peran disposisi. Ini memberi kita penjelasan terakhir untuk kesenjangan normatif/deskriptif – irasionalitas sistematis. Jika kasusnya bahwa individu berbeda-beda dalam hal seberapa besar usaha yang ingin kita lakukan untuk suatu tugas, dan bahwa jumlah upaya yang kita lakukan menentukan, sebagian, kualitas penalaran kita, maka dapat dikatakan bahwa kita tidak rasional ketika kita tidak memilih untuk menempatkan semua sumber daya kognitif kita ke dalam menghitung respons yang akan menguntungkan kita.

Stanovich (1999) menunjukkan bahwa tidak ada alasan yang diberikan oleh Panglossians atau Apologis untuk kesenjangan normatif/deskriptif yang cukup. Kesalahan kinerja acak, penerapan norma yang salah, konstruksi tugas alternatif, dan keterbatasan komputasi semuanya terbukti tidak cukup atau tidak dapat diterapkan untuk sebagian besar tugas di mana bias ditampilkan. Ini hanya menyisakan penjelasan Meliorist – irasionalitas manusia yang sistematis. Bahkan ketika kita mampu bernalar secara normatif, kita tidak selalu melakukannya. Meskipun ini mungkin tampak sebagai pandangan yang sangat negatif tentang rasionalitas manusia, ini membuka kemungkinan optimis bahwa kita dapat melakukan yang lebih baik. Ada kesenjangan nyata antara model deskriptif perilaku dan model preskriptif/normatif, setidaknya untuk sebagian besar individu pada sebagian besar tugas. Maka, mungkin saja studi matematika menjadi salah satu cara untuk mengurangi kesenjangan tersebut.

Untuk merujuk kembali ke masalah asli yang ditetapkan dalam bagian ini, penalaran 'lebih baik' ketika disebutkan dalam penelitian ini harus diartikan penalaran yang lebih dekat dengan model normatif (dan preskriptif), sejalan dengan posisi Meliorist. Dalam penalaran kondisional, interpretasi material dapat dianggap 'terbaik' karena merupakan model normatif, dianggap benar oleh ahli logika. Dari ketiga interpretasi lainnya, kondisi cacat dapat dianggap terbaik kedua karena meskipun MT tidak diterima ketika seharusnya (menurut interpretasi material) kesimpulan yang tidak valid juga tidak diterima. Di bawah interpretasi bikondisional, semua kesimpulan diterima, dua valid dan dua tidak valid, dan di bawah interpretasi konjungtif MP dan AC adalah satu-satunya yang diterima, satu valid dan satu tidak valid. Oleh karena itu interpretasi bikondisional dan konjungtif dapat dianggap sama dan 'lebih buruk' daripada interpretasi yang cacat.

### 1.3.3 Ringkasan

- Logika kondisional ('jika, maka') adalah aspek sentral dari penalaran deduktif.
- Manusia menunjukkan beragam heuristik dalam penalaran, termasuk dengan masalah kondisional.
- Data perbedaan individu Stanovich (2000) menunjukkan bahwa perspektif Meliorist tentang rasionalitas adalah akurat – secara keseluruhan, model penalaran normatif standar adalah model kompetensi yang sesuai dan karena manusia tidak selalu menalar standar ini, mereka dengan demikian menunjukkan irasionalitas sistematis.
- Ketika disebutkan dalam penelitian, penalaran yang 'lebih baik' harus diartikan sebagai penalaran yang lebih dekat dengan model normatif yang relevan.

#### 1.4 SEJARAH SINGKAT TEORI DISIPLIN FORMAL

Selama ribuan tahun telah diasumsikan bahwa orang dapat diajarkan keterampilan bernalar. Plato adalah pemegang keyakinan ini:

*"Mereka yang memiliki bakat alami untuk berhitung umumnya cepat dalam segala jenis pengetahuan; dan bahkan yang membosankan, jika mereka memiliki pelatihan aritmatika [...] menjadi jauh lebih cepat daripada yang seharusnya [...] Kita harus berusaha untuk membujuk mereka yang akan menjadi orang utama negara kita untuk pergi dan belajar aritmatika."*

*Plato (375B.C/2003, hal. 256)*

Nisbett (2009) memberikan sejarah singkat Teori Disiplin Formal, menjelaskan bagaimana kurikulum diperluas dari aritmatika Plato untuk memasukkan tata bahasa, logika, Latin dan Yunani sepanjang periode Romawi, abad pertengahan dan Renaisans. Kurikulum ini berlangsung selama berabad-abad dan akhirnya menghasilkan sistem sekolah bahasa abad kesembilan belas.

Tantangan pertama untuk TFD datang pada awal abad kedua puluh dengan perkembangan psikologi sebagai disiplin akademis. Nisbett merangkul penolakan keras terhadap TFD dengan baik dengan menyatakan bahwa "William James menertawakan gagasan bahwa pikiran memiliki otot yang dapat dilatih dengan aritmatika atau Latin". Gerakan *behavioris* menduga bahwa belajar sangat terbatas pada hubungan stimulus-respons, dipelajari melalui mekanisme seperti penguatan perilaku positif dan negatif. Jadi ide keterampilan transfer dari satu subjek ke subjek lain atau bernalar secara keseluruhan adalah kontradiksi yang lengkap. Salah satu penulis pertama yang memperdebatkan pandangan khusus domain ekstrem *behavioris* tentang pembelajaran adalah Piaget.

Teori perkembangan kognitif *Piaget* menyarankan agar kita mengembangkan keterampilan dan pengetahuan domain umum melalui empat. Tahap sensorimotor terjadi antara kelahiran dan usia 2 tahun dan ditandai dengan bayi belajar tentang dunia fisik melalui interaksi dan pengalaman sensorik mereka. Tahap praoperasional terjadi antara usia sekitar 2 dan 7 tahun, dan ditandai dengan operasi mental yang jarang dan tidak memadai secara logis. Anak dapat merepresentasikan objek melalui kata-kata dan gambar tetapi tidak dapat melakukan operasi mental yang kompleks dan tidak dapat mengambil sudut pandang orang lain.

Tahap operasional konkret terjadi antara usia 7 dan 11 tahun dan terdiri dari penggunaan penalaran logis yang tepat. Anak-anak mampu mengurutkan objek berdasarkan kategori, memahami hubungan antar objek, dan memahami sudut pandang orang lain, untuk menyebutkan beberapa contoh. Pada tahap akhir, tahap operasional formal, individu dari sekitar 11 tahun hingga dewasa mampu bernalar secara abstrak maupun konkret. Ini berarti ada pengembangan lebih lanjut dari penalaran logis, seperti dengan situasi hipotetis. Piaget percaya bahwa hanya pemikir operasional formal yang mampu membedakan kondisi material dari kondisional.

Setiap tahap yang diusulkan Piaget adalah domain-umum, dan keterampilan logis yang dikembangkan dapat diterapkan di seluruh konteks. Inhelder dan Piaget (1958) menyoroti keyakinan ini dengan menyatakan bahwa "penalaran tidak lebih dari kalkulus proposisional itu sendiri".

Pada 1950-an dan 1960-an muncul revolusi kognitif, yang menentang *behaviorisme* bahwa sebenarnya adalah mungkin dan berguna untuk menyimpulkan proses mental

daripada hanya perilaku yang dihasilkan darinya. Ini berarti ada pergeseran ke arah menyelidiki proses di balik penalaran. Apa itu juga berarti bahwa ada ruang untuk menyelidiki apa yang membuat penalaran lebih baik atau lebih buruk.

Menurut teori perkembangan kognitif Piaget, setiap orang melewati tahapan yang sama untuk memperoleh keterampilan yang sama, dan sekali lagi, menurut behaviorisme, setiap orang akan mempelajari perilaku penalaran yang sama dari stimulus yang sama, sehingga hanya ada sedikit ruang untuk variasi dalam kualitas penalaran antara individu atau untuk strategi untuk meningkatkan penalaran di luar tahapan standar atau rangsangan. Apa yang diizinkan oleh kognitivisme adalah kemungkinan bahwa ada berbagai jenis proses yang digunakan dalam penalaran, bahwa beberapa dari proses ini lebih efektif daripada yang lain, dan bahwa dengan menyelidiki proses ini dimungkinkan untuk mengubah penalaran individu.

Aturan: Jika ada A di satu sisi kartu, ada 3 di sisi lain.



**Gambar 1.3:** Tugas Pemilihan Wason

Wason adalah salah satu psikolog pertama yang menunjukkan betapa cacatnya penalaran manusia. Dia membantah gagasan Piaget tentang pengembangan penalaran logis formal dengan menunjukkan secara eksperimental bahwa, secara keseluruhan, manusia tidak terlalu logis (Wason, 1968). Dia juga berusaha untuk menjelaskan kesalahan kita dengan mengacu pada proses kognitif yang bertanggung jawab untuk menyelesaikan tugas-tugas penalaran (Johnson-Laird & Wason, 1970; Wason, 1966). Mungkin tugas yang paling terkenal dalam literatur penalaran adalah tugas pemilihan empat kartu Wason (Wason, 1966, 1968, lihat Gambar 1.3.) Peserta ditanya kartu mana yang harus mereka balikkan untuk memeriksa apakah aturan 'jika ada A pada satu sisi kartu, maka ada 3 di sisi lain' benar. Jawaban yang benar adalah memilih kartu A dan 7, tetapi hanya sekitar 10% peserta yang memberikan jawaban ini (Wason, 1968). Mayoritas menjawab baik A saja, atau keduanya A dan 3.

Johnson-Laird & Wason (1970) mengusulkan model rinci dari proses kognitif yang terlibat dalam menangani tugas seleksi, termasuk beberapa penjelasan di mana orang salah. Sebagai contoh, mereka menyarankan bahwa para pemikir awalnya hanya fokus pada item yang disebutkan dalam aturan, A dan 3, dan menganggap item yang tidak disebutkan tidak relevan. Ini mengarah pada respons yang umum dan salah 'A dan 3'. Mereka juga mengusulkan bahwa subjek tidak memiliki wawasan; wawasan parsial, atau wawasan lengkap, dan ini menentukan apakah mereka akan memilih, masing-masing, hanya intuisi pertama mereka (A dan 3), memilih kartu yang dapat memalsukan aturan juga (A, 3 dan 7) atau pilih hanya kartu yang dapat memalsukan aturan (A dan 7). Ini adalah tingkat teori tentang proses kognitif dan perbedaan individu yang tidak akan terjadi sebelum revolusi kognitif. Dengan demikian, jenis pekerjaan ini merupakan langkah maju yang signifikan dalam pemahaman kita tentang penalaran logis manusia.

Mengikuti karya terobosan Wason, psikolog lain seperti Philip Johnson-Laird, Jonathan Evans, Keith Stanovich, Daniel Kahneman dan Amos Tversky telah menggambarkan berbagai bias lain dalam penalaran manusia dan memberikan teori tentang pemrosesan kognitif yang mendasari respons yang berhasil dan yang cacat. Teori saat ini tentang penalaran dijelaskan di bawah ini dalam Bab 1.6. Apa yang penting di sini adalah teori-teori saat ini memungkinkan perbedaan dalam pemrosesan masalah penalaran, dan karenanya kemungkinan untuk meningkatkan penalaran.

## **1.5 PENELITIAN KETERAMPILAN BERNALAR DAN HUBUNGANNYA DENGAN MATEMATIKA**

TFD membuat dua klaim:

- 1) keterampilan bernalar dapat ditransfer lintas konteks,
- 2) studi matematika meningkatkan keterampilan ini ke tingkat yang lebih besar daripada disiplin ilmu lainnya.

Penelitian yang terkait dengan masing-masing asumsi tersebut akan diulas secara terpisah, diawali dengan kemampuan transfer kemampuan bernalar.

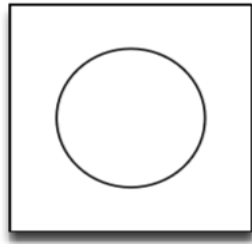
### **1.5.1 Apakah keterampilan bernalar dapat ditransfer lintas konteks?**

Salah satu tes pertama dari TFD diterbitkan oleh Thorndike dan Woodworth (1901), ketika psikolog perilaku mulai meragukan generalitas keterampilan yang dipelajari, menganjurkan hubungan stimulus-respons yang spesifik. Thorndike & Woodworth melaporkan percobaan di mana peserta dilatih dalam memperkirakan luas persegi panjang, dan peningkatan dalam memperkirakan luas bentuk lain diukur.

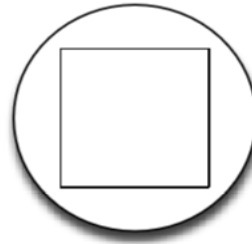
Hasil penelitian menunjukkan bahwa peningkatan dalam memperkirakan luas persegi panjang tidak diparalelkan dengan peningkatan dalam memperkirakan luas berbagai bentuk. Thorndike dan Woodworth (1901) menyarankan bahwa estimasi area adalah sekelompok fungsi daripada fungsi tunggal, dengan data spesifik (misalnya bentuk dan ukuran) menentukan fungsi mana yang diperlukan. Dalam pandangan ini, fungsi (atau keterampilan) tidak dapat dipindahtangankan karena mereka terkait erat dengan stimulus, dan setiap perubahan pada stimulus membuat fungsi menjadi berlebihan.

Kesimpulan ini didukung dalam studi skala besar dan lebih baru yang diterbitkan oleh Owen et al (2010). Dalam studi tersebut, 11.430 pemirsa program BBC 'Bang Goes The Theory' menerima pelatihan online enam minggu dalam penalaran, memori, perencanaan, keterampilan visuospasial, dan perhatian. Kemampuan penalaran sebelum dan sesudah intervensi dinilai dengan tugas tata bahasa yang dipercepat, di mana peserta membaca pernyataan tentang sepasang bentuk dan menilai apakah pernyataan itu benar. Misalnya, ketika ditunjukkan lingkaran kecil di dalam kotak yang lebih besar, peserta menilai apakah kalimat 'Lingkaran tidak lebih kecil dari kotak' itu benar.

a) Lingkaran tidak lebih kecil  
dari persegi



b) Persegi dalam lingkaran



**Gambar 1.4:** Contoh tugas penalaran yang digunakan oleh Owen et al (2010) menunjukkan (a) kalimat salah dan (b) benar tentang pasangan bentuk.

Selama intervensi peserta berlatih tiga tugas penalaran yang berbeda, serta memori, perencanaan dan tugas visuospasial, dengan umpan balik. Dalam satu tugas penalaran, peserta diminta untuk menyimpulkan hubungan berat dari gambar jungkat-jungkit dengan objek di setiap ujungnya untuk memilih objek terberat dari tiga pilihan yang disajikan secara terpisah. Pada tugas penalaran kedua, peserta diminta untuk memilih satu dari empat bentuk yang bervariasi dalam bentuk, warna, dan isi. Tugas penalaran terakhir mengharuskan peserta untuk memindahkan peti dari tumpukan dengan mengacu pada pola peti secara keseluruhan. Meskipun ada peningkatan dalam setiap tugas yang dilatih, tidak ada bukti transfer ke tugas penalaran gramatikal yang tidak terlatih, yang oleh penulis dianggap terkait erat. Ini tampaknya menjadi bukti bahwa keterampilan bernalar terkait erat dengan konteks, seperti yang diusulkan Thorndike dan Woodworth (1901).

Dalam gagasan pelatihan keterampilan bernalar, Cheng, Holyoak, Nisbett dan Oliver (1986) menemukan bahwa bahkan satu semester saja dalam logika standar tidak mengarah pada peningkatan kinerja pada empat masalah Tugas Seleksi Wason. Kursus ini terdiri dari sekitar 40 jam pengajaran dan mencakup modus ponens, modus tollens, penolakan anteseden, penegasan konsekuensi dan perbedaan antara pernyataan kondisional dan kondisional materi. Mahasiswa juga diajarkan dengan kalimat kontekstual dan tabel kebenaran formal. Tampaknya masuk akal untuk mengharapkan bahwa setelah pelatihan tersebut peserta harus sangat kompeten dalam menangani masalah penalaran kondisional, dan bahkan sulit untuk membayangkan cara yang lebih menjanjikan untuk meningkatkan kompetensi logis mahasiswa. Meskipun demikian, ada penurunan kesalahan yang tidak signifikan hanya 3%. Namun, ini bisa jadi karena ukuran yang digunakan. Baru-baru ini disarankan bahwa Tugas Seleksi mungkin tidak benar-benar mengukur kemampuan penalaran kondisional, terutama versi tematik (terletak dalam konteks dunia nyata).

Sperber dkk. (1995) telah menyarankan bahwa kinerja Seleksi Tugas sangat dipengaruhi oleh mekanisme relevansi yang mendahului mekanisme penalaran apa pun. Ketika dihadapkan dengan masalah penalaran, atau teks lain, kita perlu memahami makna yang dimaksudkan oleh penulis. Dalam kasus tugas seleksi, peserta diminta untuk menilai relevansi masing-masing kartu dengan aturan sehingga penilaian relevansi yang berasal dari proses pemahaman memberikan jawaban intuitif untuk masalah dan tidak ada kebutuhan eksplisit untuk terlibat dalam alasan lebih lanjut. Ini mungkin sumber dari bias pencocokan yang meresap yang dijelaskan di atas. Akun ini telah didukung dalam beberapa penelitian dan

ini menyiratkan bahwa Tugas Seleksi tidak benar-benar mengukur proses penalaran sama sekali. Argumen Sperber diuraikan lebih lanjut dalam Bab 3.

Dalam studi serupa tentang pengajaran penalaran, White (1936) menyelidiki efek pelatihan logika pada kemampuan penalaran anak laki-laki berusia 12 tahun. Satu kelas menghabiskan satu jam per minggu selama tiga bulan untuk diajarkan logika, termasuk deduksi, induksi dan silogisme, sementara kelas lain tidak. Pada akhir tiga bulan mahasiswa diberikan tes penalaran yang diukur antara lain validasi silogisme. Kelas yang diajarkan logika mendapat nilai yang lebih tinggi pada tes penalaran dibandingkan kelas kontrol. Para penulis menyimpulkan, sebaliknya untuk Cheng et al. (1986), bahwa bernalar logis dapat diajarkan. Mungkin perbedaan antara temuan Cheng et al. (1986) dan White (1936) karena usia partisipan, bisa jadi strategi penalaran anak-anak lebih mudah dibentuk daripada orang dewasa. Kemungkinan lain adalah ukuran yang digunakan oleh Cheng et al. (1986) tidak sesuai untuk mengukur perkembangan, sebagaimana diuraikan di atas.

Lebih lanjut pendukung gagasan bahwa penalaran dapat diajarkan, Lehman, Lempert dan Nisbett (1988) menemukan bahwa sejauh mana penalaran kondisional ditingkatkan pada mahasiswa pascasarjana adalah fungsi dari studi gelar mereka. Mahasiswa dalam psikologi, kedokteran, hukum dan kimia diuji pada penalaran kondisional, penalaran statistik dan metodologis, dan penalaran verbal pada tahun pertama dan ketiga studi mereka. Mahasiswa kedokteran meningkat dalam penalaran verbal dan baik psikologi dan mahasiswa kedokteran meningkat dalam penalaran statistik dan metodologis. Yang paling relevan di sini adalah mahasiswa psikologi, hukum, dan kedokteran semuanya meningkat dalam penalaran kondisional sebagaimana diuji oleh satu abstrak, satu tugas seleksi berbingkai kausal, dan satu tugas seleksi berbingkai izin dan satu tugas seleksi bikondisional. Namun, seperti disebutkan di atas, tidak jelas keterampilan apa yang sebenarnya diukur dalam Tugas Seleksi kontekstual. Oleh karena itu, temuan harus diperlakukan dengan hati-hati.

Kosonen dan Winne (1995) menemukan bukti transfer setelah peserta pelatihan dalam penalaran statistik. Pelatihan 'hukum bilangan besar' (yang menyatakan bahwa semakin besar ukuran sampel, semakin representatif) meningkatkan kemampuan mahasiswa untuk bernalar dengan masalah sehari-hari yang membutuhkan hukum bilangan besar dalam berbagai domain. Sebelum pelatihan, peserta lebih mungkin untuk menalar secara statistik tentang masalah yang secara eksplisit mengacu pada probabilitas daripada yang tidak, tetapi ini sama dengan setelah pelatihan.

Karena para peserta pascasarjana dan mahasiswa menggunakan aturan statistik sebelum instruksi, Kosonen dan Winne (1995) menyarankan bahwa aturan penalaran informal dipelajari melalui pengalaman dan bahwa ini direkrut dan disempurnakan dengan pelatihan yang diberikan, membuat mereka dipindahtangankan. Ini juga merupakan kesimpulan dari Fong, Krantz dan Nisbett (1986), yang menemukan bahwa pelatihan dalam hukum bilangan besar meningkatkan kinerja peserta bahkan ketika mereka diuji dengan konteks masalah yang sama sekali berbeda dengan pelatihan. Lebih lanjut, Fong dan Nisbett (1991) menemukan bahwa transfer masih terbukti setelah penundaan dua minggu.

Untuk menguji keumuman bernalar kritis, S´a et al. (1999) mengukur bias keyakinan partisipan dalam tugas verbal (silogisme) dan tugas non-verbal (penilaian tinggi badan). Tugas silogisme terdiri dari 24 silogisme dengan konten dunia nyata, seperti 'semua makhluk hidup membutuhkan air; mawar membutuhkan air; oleh karena itu, mawar adalah makhluk hidup' (lihat Gambar 1.5). Dalam tugas penilaian tinggi badan, peserta diperlihatkan foto dengan

ukuran yang besar dan persis dengan ukuran pria dan wanita yang duduk dan diminta untuk memperkirakan tinggi badan mereka. Dalam versi ekologi, ketinggian yang ditampilkan mencerminkan perbedaan tinggi pria-wanita yang sebenarnya. Dalam versi yang cocok, gambar yang ditampilkan adalah pria dan wanita dengan tinggi badan yang cocok, dan peserta diberitahu untuk mengabaikan gender dalam membuat penilaian.

Dalam setiap tugas, tanggapan peserta dapat dipengaruhi oleh keyakinan sebelumnya dan/atau validitas logis. Kepercayaan adalah sejauh mana penilaian dipengaruhi oleh keyakinan sebelumnya atas validitas ketika keduanya berada dalam konflik. Para penulis menemukan korelasi yang signifikan ( $r=.209$ ) antara bias keyakinan dalam tugas silogisme dan bias keyakinan dalam tugas penilaian tinggi yang cocok dan ada korelasi yang lebih besar lagi antara bias keyakinan dalam dua versi tugas penilaian tinggi ( $r= 0,526$ ). Sementara korelasi yang signifikan bertentangan dengan pandangan spesifisitas domain yang ekstrim, seperti dilakukan Thorndike & Woodworth (1901), mereka mendukung gagasan bahwa generalitas terkait dengan kesamaan antara rangsangan tugas; korelasinya lebih kuat antara dua tugas penilaian tinggi badan non-verbal daripada antara keduanya dan tugas silogisme verbal.

Dapat dipercaya, Valid :	Dapat dipercaya, Valid :
Premis : Semua ikan dapat berenang Tuna adalah ikan	Premis : Semua yang hidup membutuhkan air mawar membutuhkan air
Kesimpulan : Tuna dapat berenang	Kesimpulan : Mawar adalah benda hidup
Tidak dapat dipercaya, valid :	Tidak dapat dipercaya, valid :
Premis : Semua yang berkaki empat berbahaya Anjing tidak berbahaya	Premis : Semua senapan berbahaya Ular berbahaya
Kesimpulan : Anjing tidak memiliki empat kaki	Kesimpulan : Ular merupakan senapan
Netral, valid :	Netral, valid :
Premis : Semua Biskuit rasanya enak. Oreo adalah ramadion.	Premis : Semua manusia mengenakan pakaian. Andi mengenakan pakaian
Kesimpulan : Oreo rasanya enak.	Kesimpulan : Andi adalah manusia

**Gambar 1.5:** Contoh setiap jenis item dari tugas Belief Bias Silogisme.

Pada tema yang sama, Toplak dan Stanovich (2002) menyelidiki keumuman domain dari penalaran disjungtif ('baik, atau'). Mereka memberi peserta sembilan tugas yang semuanya bergantung pada pertimbangan menyeluruh dari semua kemungkinan keadaan dunia yang diberikan aturan 'baik, atau' (seperti masalah Ksatria dan Knaves yang ditunjukkan pada Gambar 1.6). Jika ini adalah keterampilan tunggal yang digeneralisasi untuk semua tugas, seharusnya ada korelasi besar antara kinerjanya. Bahkan, kinerja pada tugas-tugas menunjukkan kekhususan. Hal ini menyebabkan penulis untuk berkonsentrasi pada lima tugas yang tak terbantahkan membutuhkan penalaran disjungtif. Di antara tugas-tugas ini, lima dari sepuluh korelasi secara statistik signifikan, masih menunjukkan tingkat spesifisitas domain yang cukup besar antara tugas-tugas yang secara yakin dianggap menggunakan keterampilan yang sama, meskipun dalam konteks yang berbeda. Sekali lagi, ini tampaknya mendukung argumen Thorndike dan Woodworth (1901) bahwa setiap perubahan rangsangan mencegah keterampilan yang cocok untuk konteks lain diterapkan dengan benar.

Ada dua studi terakhir yang dapat menginformasikan pertanyaan apakah keterampilan bernalar dapat dialihkan, meskipun keduanya memiliki hambatan signifikan yang berarti mereka tidak boleh dianggap definitif. Sanz de Acedo Lizarraga, Sanz de Acedo Baquedano dan Soria Oliver (2010) melihat pengaruh intervensi pengajaran selama satu tahun terhadap



keterampilan bernalar di dua sekolah Spanyol dan Lehmann (1963) melihat perkembangan bernalar kritis selama empat tahun pendidikan perguruan tinggi .

Dalam Sanz de Acedo Lizarraga dkk. (2010) penelitian, partisipan berusia 11 sampai 13 tahun dan berasal dari dua sekolah di Spanyol. Satu sekolah secara acak ditugaskan ke kondisi eksperimental dan satu ke kondisi kontrol. Kondisi eksperimental melibatkan metode pengajaran 'Bernalar Aktif dalam Konteks Akademik', yang dirancang untuk meningkatkan keterampilan bernalar pada mahasiswa. Intervensi berlangsung selama satu tahun akademik dan dilaksanakan selama dua belas jam per minggu. Peserta menyelesaikan tes pra dan pasca intervensi yang mengukur berbagai kemampuan kognitif.

A : B adalah seorang penjahat  
 B : A dan C adalah tipe yang sama  
 Apakah identitas C ?

**Gambar 2.6:** Masalah Knights and Knaves disjungtif yang digunakan oleh Toplak dan Stanovich (2002), diadaptasi dari Shafir (1994).

Intervensi terbukti meningkatkan penalaran verbal, abstrak dan numerik, kreativitas, dan prestasi akademik ke tingkat yang lebih besar daripada metode pengajaran konvensional yang digunakan dengan kelompok kontrol. Meskipun ini mendukung keterampilan bernalar secara umum, ada kelemahan penting pada desain penelitian. Intervensi ditugaskan secara acak di tingkat sekolah daripada tingkat peserta, yang berarti perbedaan antara sekolah yang mungkin telah mempengaruhi hasil, seperti kemampuan umum mahasiswa, tidak dikontrol. Itu juga tidak ditugaskan secara membabi buta. Masalah ini sangat penting karena penelitian ini melibatkan metode pengajaran yang sepenuhnya baru, baik untuk guru maupun mahasiswa, yang menurut mereka akan memiliki manfaat yang diharapkan. Hal ini akan menyebabkan peningkatan minat, antusiasme, dan harapan dari mereka yang terlibat, dan apa yang disebut efek Hawthorne<sup>1</sup> ini mungkin yang sebenarnya bertanggung jawab atas efek yang diamati.

Studi Lehmann (1963) menyelidiki perubahan dalam penalaran kritis mahasiswa Indonesia dan keyakinan stereotip selama studi pendidikan tinggi mereka. Antara tahun pertama dan keempat mereka, ia menemukan peningkatan yang signifikan dalam bernalar kritis, seperti yang diukur oleh Tes Pendidikan Kritis oleh Pemerintah Indonesia, dan penurunan keyakinan stereotip (yakni, para peserta menjadi lebih fleksibel dan kurang otoriter). Para mahasiswa tidak secara khusus dilatih untuk efek-efek ini, jadi mungkin saja perubahan-perubahan itu muncul sebagai suatu generalisasi dari sesuatu dari pendidikan mereka. Atau, karena tidak ada kelompok kontrol, mungkin terjadi perubahan ini terjadi pada semua orang usia kuliah terlepas dari apakah mereka sedang menempuh pendidikan atau tidak. Namun demikian, kedua studi pendidikan ini memberikan beberapa harapan untuk bernalar kritis secara umum.

---

<sup>1</sup> Efek Hawthorne mengacu pada pengamatan bahwa peserta dalam studi penelitian dapat mengubah perilaku mereka hanya sebagai hasil dari yang diamati, terlepas dari manipulasi. Ini pertama kali diidentifikasi oleh Henry Landsberger yang melakukan studi di pabrik Hawthorne Works tentang pengaruh tingkat pencahayaan terhadap produktivitas. Peningkatan produktivitas terjadi saat penelitian sedang dilakukan dan penurunan terjadi saat penelitian dihentikan. Disarankan bahwa para pekerja menjadi lebih produktif hanya karena mereka sedang diamati (Landsberger, 1958).

Singkatnya, bukti untuk klaim pertama yang dibuat oleh TFD, keterampilan bernalar dapat ditransfer lintas konteks, tidak konsisten dan oleh karena itu cukup lemah. Sementara beberapa penelitian telah menemukan beberapa tingkat umum dari keterampilan bernalar, yang lain telah menemukan kekhususan yang lengkap.

Mungkin ini karena setiap studi telah melihat aspek penalaran yang berbeda dan mungkin saja setiap keterampilan memiliki posisi yang berbeda pada kontinum kekhususan dan umum. Penalaran statistik dan kerentanan terhadap keyakinan (S'a et al., 1999) akan tampak lebih jauh ke arah akhir spektrum umum daripada penalaran disjungtif (Toplak & Stanovich, 2002) dan penalaran tentang bentuk daerah (Thorndike & Woodworth, 1901). Berbagai macam diajarkan dalam metode pengajaran 'Bernalar Aktif dalam Konteks Akademik' dan di pendidikan tinggi, jadi tidak jelas secara pasti keterampilan apa yang digeneralisasikan dalam kasus Sanz de Acedo Lizarraga et al. (2010) dan Lehmann (1963).

Penafsiran inkonsistensi ini sejalan dengan saran Nisbett dan rekan-rekannya bahwa kesulitan mengajar jenis penalaran tergantung pada sejauh mana orang sudah memiliki pemahaman intuitif aturan (Nisbett, Fong, Lehman & Cheng, 1987; Nisbet, 2009). Seperti yang juga disarankan oleh Kosonen dan Winne (1995), orang mungkin memiliki pemahaman dasar tentang hukum statistik yang dilantik melalui pengalaman, yang membuat mereka tersedia untuk pelatihan. Mungkin penalaran disjungtif dan geometris (area bentuk) belum berada dalam repertoar kognitif aturan dasar dan karena itu lebih sulit untuk diajarkan dengan cara yang dapat ditransfer.

### 1.5.2 Apakah studi matematika meningkatkan 'keterampilan bernalar'?

Pindah ke asumsi kedua dari TFD, bahwa matematika meningkatkan keterampilan bernalar ke tingkat yang lebih besar daripada mata pelajaran lain, buktinya lagi beragam. Seperti asumsi pertama, Thorndike (1924) menerbitkan salah satu studi pertama yang menguji pengaruh mata pelajaran sekolah pada kemampuan penalaran.

Thorndike (1924) menggunakan desain pre-test/intervention/post-test, di mana intervensinya adalah satu tahun pendidikan sekolah dengan mata pelajaran yang bervariasi dengan tes awal dan akhir mengukur kecerdasan umum (kemampuan tes berbagai mata pelajaran tipe sekolah, seperti aritmatika dan analogi kata dan lawan kata). Dia menemukan bahwa mata pelajaran yang diambil oleh mahasiswa hanya memiliki pengaruh minimal pada skor pada tes kecerdasan umum. Bahasa, kimia, dan trigonometri dikaitkan dengan peningkatan terbesar, meskipun kecil, sementara aritmatika, geometri, dan aljabar dikaitkan dengan peningkatan hampir di atas nol. Jadi sekali lagi, penelitian Thorndike melukiskan gambaran suram untuk TFD.



Anda hanya boleh membalik label yang perlu Anda periksa untuk memastikan labelnya benar.

**Gambar 1.7:** Masalah kondisional skema kausal yang digunakan dalam Lehman dan Nisbett (1990).

Namun, Lehman dan Nisbett (1990) menemukan beberapa dukungan untuk versi TFD. Mereka menguji pada penalaran statistik dan metodologis, penalaran kondisional dan penalaran verbal mahasiswa sarjana Indonesia di tahun pertama dan keempat mereka. Mereka membandingkan mahasiswa yang belajar ilmu alam, humaniora, ilmu sosial, dan psikologi. Meskipun tidak ada perbedaan antara kelompok pada salah satu ukuran pada tes pertama atau pada skor SAT, ada beberapa efek posttest dari disiplin yang dipelajari. Mahasiswa ilmu sosial dan psikologi meningkat secara dramatis pada penalaran statistik dan metodologis, sedangkan mahasiswa ilmu alam dan humaniora meningkat secara signifikan pada penalaran kondisional, menjadi lebih banyak materi pada satu abstrak, satu kerangka kausal dan satu tugas Seleksi berbingkai izin, dan satu Tugas Seleksi bikondisional.

Yang paling penting, ada korelasi antara perubahan penalaran kondisional materi dan jumlah mata kuliah matematika yang diambil (Lehman & Nisbett, 1990). Seperti yang telah dinyatakan, penalaran kondisional merupakan komponen penting dari penalaran logis (Anderson & Belnap, 1975; Braine, 1978; Inglis & Simpson, 2008). Itu termasuk dalam braket penalaran deduktif, yang Polya (1954) anggap sangat penting dalam pembuktian matematis. Korelasi dengan jumlah modul matematika yang diambil cukup signifikan untuk semua jurusan ( $r=0,31$ ), tetapi lebih kuat lagi jika berfokus pada jurusan IPA ( $r=0,66$ ), yang mengambil sebagian besar mata kuliah matematika.

Penulis menyimpulkan bahwa penalaran dapat diajarkan, dan bahwa disiplin yang berbeda mengajarkan jenis penalaran yang berbeda (Lehman & Nisbett, 1990). Ini sangat menjanjikan bagi TFD, terutama mengingat betapa pentingnya beberapa temuan yang dibahas sampai saat ini. Namun, seperti yang disebutkan di atas dalam pembahasan Lehman et al. (1988), tidak jelas keterampilan apa yang sebenarnya diukur oleh Tugas Seleksi. Sekali lagi, hasilnya harus dipertimbangkan dengan hati-hati.

Kemampuan penalaran kondisional juga diselidiki oleh Inglis dan Simpson (2008), yang membandingkan mahasiswa matematika dan seni. Mereka memberikan 32 item abstrak tugas inferensi kondisional kepada mahasiswa dan mengamati bahwa mahasiswa matematika tampil secara signifikan lebih sesuai dengan kondisi materi daripada mahasiswa seni. Sekali lagi, ini tampaknya mendukung TFD, tetapi ada dua masalah dengan penelitian ini: tidak ada ukuran kecerdasan, yang mungkin berbeda antar kelompok dan bertanggung jawab atas perbedaan skor inferensi kondisional, dan tidak ada komponen longitudinal, yang membuat tidak mungkin untuk menyimpulkan perkembangan. Masalah-masalah ini dibahas dalam studi kedua.

Inglis dan Simpson (2009a) mencocokkan sekelompok mahasiswa matematika dengan sekelompok mahasiswa perbandingan pada skor kecerdasan yang ditentukan oleh tes AH5 (Heim, 1969), dan sekali lagi memberi mereka tugas inferensi kondisional. Kecerdasan berhubungan positif dengan skor penalaran kondisional materi dan lebih tinggi pada mahasiswa matematika daripada mahasiswa non-matematika sebelum sub-kelompok yang cocok dipilih. Namun demikian, pencocokan kelompok untuk kecerdasan masih meninggalkan perbedaan yang signifikan antara mereka pada materi skor kondisional, dengan mahasiswa matematika lagi mengungguli kelompok perbandingan. Dengan kata lain, kelompok matematika lebih normatif dalam penalaran kondisional mereka melebihi dan di atas skor kecerdasan mereka yang lebih tinggi.

Para mahasiswa matematika diuji ulang pada akhir tahun pertama studi mereka untuk mencari perkembangan dalam kemampuan penalaran kondisional. Ada perubahan rata-rata

hanya 1,8% terhadap interpretasi materi, yang tidak mendekati signifikansi. Kurangnya perbaikan meninggalkan dua kemungkinan penjelasan untuk perbedaan awal antara kelompok saat masuk ke universitas: studi matematika pasca-wajib tetapi pra-universitas bertanggung jawab, yaitu A level<sup>2</sup>; atau mereka yang lebih normatif dalam penalaran kondisional mereka secara tidak proporsional disaring untuk mempelajari matematika tingkat universitas. Membedakan dua kemungkinan ini pada dasarnya adalah tujuan utama dari penelitian ini.

Dengan asumsi bahwa hasil yang ditemukan oleh Inglis dan Simpson (2008, 2009a) cukup untuk mengatakan bahwa sarjana matematika lebih normatif dalam penalaran kondisional mereka daripada sarjana lain saat masuk ke universitas, maka masalah yang harus diselesaikan bukanlah apakah matematikawan lebih baik dalam penalaran, tetapi apakah matematika membuat mereka lebih baik dalam penalaran. Peserta matematika Inglis dan Simpson mungkin telah mengungguli peserta perbandingan mereka karena belajar matematika di tingkat A mengubah kemampuan penalaran mereka untuk membawa mereka ke titik itu, seperti yang disarankan TFD, atau karena mereka selalu lebih normatif dalam penalaran mereka, terlepas dari matematika apa pun yang dipelajari, dan disaring untuk mempelajari matematika pada tingkat lanjutan karena sifat yang melekat ini. Cara untuk membedakan kemungkinan ini adalah dengan studi longitudinal yang melihat perubahan kemampuan penalaran di seluruh tingkat A (tahap pertama pendidikan pasca-wajib), di mana mahasiswa matematika dibandingkan dengan mahasiswa non-matematika. Meskipun tidak mungkin menentukan sebab akibat tanpa desain eksperimental (lihat Bab 3), hipotesis penelitian pengembangan dapat dibedakan dari hipotesis penelitian pra-disposisi.

Kesimpulannya, penelitian yang diulas di sini tidak jelas apakah asumsi TFD dibenarkan atau tidak. Meskipun tampak bahwa beberapa keterampilan bernalar tidak dapat ditransfer dari konteks asli di mana mereka dipelajari, ada beberapa bukti bahwa matematikawan beralasan lebih normatif pada Tugas Inferensi Kondisional. Untuk menguji TFD secara efektif, diperlukan studi longitudinal di mana mahasiswa matematika kelas enam dan mahasiswa universitas dibandingkan dengan mahasiswa dari disiplin lain, dan perkembangan masing-masing kelompok diukur dan dikontraskan. Hanya dengan demikian akan mungkin untuk menguji dua klaim TFD secara efektif. Dua studi oleh Inglis dan Simpson (2008, 2009), bersama dengan studi Lehman dan Nisbett (1990), memberikan garis penyelidikan yang menjanjikan dalam studi lebih lanjut: jika TFD benar, tampaknya penalaran kondisional mungkin merupakan keterampilan bernalar. bahwa matematika berkembang.

### ***Ringkasan bagian***

- Pertanyaan apakah keterampilan bernalar dapat ditransfer lintas konteks atau domain telah diperdebatkan dengan hangat dan ada bukti yang mendukung dan menentangnya.
- Ada beberapa bukti bahwa belajar matematika mungkin terkait dengan kemampuan penalaran kondisional (Inglis & Simpson, 2008) tetapi tidak selalu peningkatan di

---

<sup>2</sup> Tingkat A adalah kursus dua tahun opsional yang diambil setelah wajib belajar. Mahasiswa biasanya mengambil empat level A dalam mata pelajaran yang dipilih dan hasilnya digunakan oleh universitas untuk memilih sarjana yang masuk. Tahun pertama level A disebut level Advanced Subsidiary (AS), dan merupakan kualifikasi tersendiri.

dalamnya (Inglis & Simpson, 2009a) atau peningkatan dalam penalaran statistik, metodologis atau verbal (Lehman & Nisbett, 1990) .

## 1.6 PSIKOLOGI PENALARAN

TFD menyarankan bahwa mempelajari matematika meningkatkan penalaran, tetapi tidak menguraikan perubahan kognitif apa yang sebenarnya terjadi yang dapat menyebabkan hal ini. Namun, sejumlah besar penelitian telah dilakukan ke dalam psikologi penalaran secara independen dari hubungan apa pun dengan matematika, dan saat ini ada beberapa teori dominan yang dapat membantu mengidentifikasi mekanisme yang mungkin. Teori-teori tersebut ditinjau dalam tiga bagian tergantung pada berapa banyak sistem penalaran atau tipe proses yang mereka ajukan: teori proses tunggal, proses ganda, dan tiga proses.

### 1.6.1 Teori proses tunggal

Perlu dicatat di sini bahwa teori proses tunggal tidak selalu bersaing satu sama lain atau dengan teori proses ganda. Setidaknya beberapa teori proses tunggal hanya mencoba untuk menjelaskan jenis penalaran tertentu daripada penalaran secara keseluruhan, dan mereka juga dapat dianggap sebagai penjelasan dari salah satu jenis proses dalam teori proses ganda. Ini harus menjadi lebih jelas karena masing-masing teori dibahas di bawah ini.

### 1.6.2 Teori Model Mental.

Model mental teori penalaran (Johnson-Laird & Byrne, 1991; Johnson-Laird & Byrne, 2002; Johnson-Laird, 2008) menunjukkan bahwa peserta membuat dalam pikiran mereka model tempat tugas seperti yang mereka pahami, dan mereka beralasan dari model tugas mereka. Model mental mewakili kemungkinan keadaan dunia yang diberikan bahwa premis mayor adalah benar. Misalnya, ketika diberikan pernyataan kondisional 'jika hujan maka saya membawa payung', model awalnya mungkin:

*Hujan*            *bawa payung*

*Lapar*            *Makan*

*Lelah*            *Istirahat*

...

Baris pertama menunjukkan model di mana 'hujan' dan 'membawa payung' adalah benar. Elipsis menunjukkan model implisit yang mengingatkan penalaran bahwa mungkin tidak benar, tanpa benar-benar membangun model secara eksplisit. Hal ini terjadi karena membangun model mental membutuhkan memori kerja yang terbatas.

**Tabel 1.4** Tabel kebenaran untuk 'jika hujan maka saya membawa payung' di mana t = benar dan f = salah, dengan asumsi bahan kondisional.

Hujan	Bawa Payung	Jika hujan maka saya bawa payung
t	t	t
t	f	f
f	t	t
f	f	f

Setiap model mental mewakili baris 'benar' dalam tabel kebenaran, meskipun seperti yang telah dicatat model mungkin tidak mewakili semua informasi di baris yang sesuai, dan mungkin ada baris yang tidak diwakili secara eksplisit sama sekali (Johnson-Laird & Byrne ,

2002). Tabel kebenaran materi lengkap untuk pernyataan kondisional ini ditunjukkan pada Tabel 1.4.

Penalaran dapat membangun beberapa model dari premis untuk mewakili semua kemungkinan seperti yang mereka lihat, dan dari model ini mereka menarik kesimpulan yang berlaku untuk semuanya. Misalnya, seorang penalaran mungkin melihat kemungkinan sebagai:

<i>Hujan</i>	<i>Bawa payung</i>
<i>Tidak Hujan</i>	<i>Tak perlu Bawa payung</i>

...

Jika premis minor yang disajikan adalah 'membawa payung' maka berdasarkan model mental mereka, si penalar dapat menyimpulkan bahwa 'hujan' harus mengikuti, karena tidak ada model di mana kesimpulan itu tidak benar, tetapi ini akan menjadi penegasan yang tidak valid dari konsekuensinya. deduksi.

Langkah terakhir dalam proses penalaran yang diusulkan oleh teori model mental adalah penalaran mencoba untuk memikirkan contoh tandingan untuk model dan kesimpulan mereka. Jika mereka tidak menemukan contoh tandingan, mereka menerima kesimpulan mereka, jika tidak, mereka dapat merekonstruksi model mereka untuk memasukkan contoh tandingan dan kemudian menarik kesimpulan baru.

Untuk mendemonstrasikan teori ini lebih jauh, kita dapat membayangkan bagaimana seseorang menghadapi pernyataan kondisional abstrak 'jika p maka q'. Model awal mungkin:

<i>p</i>	<i>q</i>
----------	----------

...

Baris pertama menunjukkan model di mana p dan q keduanya benar dan elipsis menunjukkan model implisit yang mengingatkan si pemberi alasan bahwa p mungkin tidak benar. Jika premis minor setelah 'jika p maka q' adalah modus ponens inferensi 'p', maka peserta dapat dengan mudah menarik kesimpulan 'q' dari model awal mereka. Demikian pula, jika premis minor adalah penegasan dari kesimpulan konsekuen 'q', maka akan mudah bagi peserta untuk salah menyimpulkan 'p'. Namun, dalam kasus modus tollens, premis awal tidak cukup untuk menarik kesimpulan. Ketika premis minor berbunyi 'bukan q', si pemikir harus mengeluarkan model implisitnya untuk menarik kesimpulan apa pun. Jika seorang reasoner benar-benar membangun semua model yang mungkin di bawah aturan kondisional, mereka akan berakhir dengan:

<i>p</i>	<i>q</i>
<i>bukan p</i>	<i>q</i>
<i>bukan p</i>	<i>bukan q</i>

Dalam hal ini, premis 'bukan q' benar dalam model terakhir, di mana 'bukan p' juga benar, jadi inilah kesimpulan yang ditarik. Para peserta kemudian akan mencari contoh tandingan untuk kesimpulan ini, sebuah model di mana kesimpulannya tidak mungkin benar. Jika mereka gagal menemukan model seperti itu, mereka menerima kesimpulannya.

Kesalahan dalam penalaran terjadi ketika peserta gagal menyempurnakan model mental implisit mereka, atau ketika mereka melewatkan contoh tandingan yang valid. Ini mungkin terjadi karena keterbatasan memori kerja dan/atau kurangnya usaha. Dengan demikian, ada cara untuk meningkatkan penalaran di bawah teori model mental: meningkatkan kapasitas memori kerja, membantu memori kerja dengan desain tugas, atau meningkatkan upaya peserta dalam penalaran mereka. Dalam hal bagaimana matematika secara khusus dapat meningkatkan kemampuan penalaran, peningkatan kapasitas memori

kerja atau perubahan disposisi bernalar (seperti peningkatan kenikmatan bernalar yang penuh usaha), mungkin dihasilkan dari belajar matematika.

### **1.6.3 Teori Logika Mental**

Teori logika mental (Rips, 1989; O'Brien, 2009; O'Brien & Manfrinati, 2010), juga dikenal sebagai teori aturan mental, teori aturan inferensi, dan teori deduksi alami, menunjukkan bahwa manusia memiliki aturan logika bawaan. sistem. Ini agak mirip dengan proposisi Piaget. Ketika kita menyelesaikan tugas penalaran, menurut teori logika mental, kita mengikuti serangkaian langkah logis untuk mencapai kesimpulan. Ini adalah kasus untuk semua jenis tugas penalaran, apakah itu kontekstual atau abstrak, akrab atau baru. Sistem ini bawaan, seperti aturan yang digunakannya. Namun, sementara kami mengikuti pola deduksi logis formal, sistem kami dan aturannya tidak sempurna dan ada banyak ruang untuk kesalahan, itulah sebabnya manusia menampilkan banyak perilaku seperti yang kami lakukan. Misalnya, kita mungkin memiliki akses bawaan ke pengurangan modus ponens tetapi bukan pengurangan modus tollens.

Menurut teori logika mental, cara di mana penalaran dapat ditingkatkan adalah dengan bekerja pada keakuratan sistem deduksi alami, mungkin meskipun studi logika formal karena strukturnya sama dengan sistem bawaan kita. Hal ini menunjukkan bahwa matematika sebenarnya dapat menjadi cara yang cukup baik untuk meningkatkan penalaran umum, karena menghadapkan mahasiswa pada logika dalam bentuk bukti matematika, dan dalam beberapa kasus bahkan mungkin mengajarkan logika secara langsung.

### **1.6.4 Teori Skema Penalaran Pragmatis**

Cheng dan Holyoak (1985) mengusulkan agar kita bernalar berdasarkan skema penalaran pragmatis. Ini adalah struktur pengetahuan konteks-sensitif yang diinduksi dari pengalaman sehari-hari. Mereka berhubungan dengan set aturan seperti izin, kewajiban, dan sebab-akibat dan mereka memungkinkan kita untuk menangani secara efektif jenis situasi penalaran yang kita hadapi sehari-hari. Misalnya, jika seseorang meminjamkan uang kepada kami, skema kewajiban kami dipanggil untuk memberi tahu kami bahwa kami berkewajiban untuk membayarnya kembali.

Aturan yang tidak terkait dengan pengalaman kita sebelumnya, termasuk aturan abstrak dan aturan yang ditetapkan dalam konteks yang tidak kita kenal, tidak akan memanggil skema dan karena itu akan sangat sulit untuk dipecahkan. Jika kita tidak memiliki skema yang relevan, kita harus bergantung pada aturan logis, yang hanya sedikit dari kita yang dikatakan kompeten (Cheng & Holyoak, 1985). Hal ini didukung oleh temuan bahwa 81% peserta dapat dengan benar menyelesaikan versi tematik Tugas Seleksi, sementara hanya 15% yang dapat menyelesaikan versi abstrak (Johnson-Laird, Legrenzi & Legrenzi, 1972). Namun, Sperber et al. (1995) menunjukkan bahwa ini tidak selalu terjadi, dan setidaknya teori skema penalaran pragmatis bukanlah teori penalaran yang komprehensif jika hanya mencoba untuk menjelaskan penalaran dalam konteks tertentu dan mengacu pada teori seperti logika mental untuk menjelaskan penalaran. dalam konteks yang tidak memunculkan skema.

Fitur penting dari skema penalaran pragmatis adalah mereka tidak selalu valid secara logis, tetapi heuristik yang membantu kita untuk menyelesaikan sebagian besar tugas penalaran yang akan kita temui dalam kehidupan sehari-hari. Misalnya, skema izin menyediakan empat aturan berikut:

- 1) Jika suatu tindakan akan diambil, maka prasyarat harus dipenuhi.
- 2) Jika suatu tindakan tidak akan diambil, prasyarat tidak perlu dipenuhi.

- 3) Jika prasyarat terpenuhi, maka tindakan dapat diambil.
- 4) Jika prasyarat tidak terpenuhi, maka tindakan tidak boleh dilakukan

Skema ini akan diperoleh dengan memberikan pernyataan kondisional 'jika seseorang minum alkohol, maka mereka harus berusia di atas 18 tahun.' Namun, aturan 3 secara logis tidak konsisten dengan aturan 1. Ini mungkin terjadi bahwa jika seseorang ingin minum alkohol, mereka harus di atas 18 tahun, tetapi tidak selalu berarti bahwa jika seseorang berusia di atas 18 tahun, mereka boleh minum alkohol. Mungkin dalam kasus mereka hamil atau akan mengendarai mobil, dalam hal ini prasyarat berusia di atas 18 tahun diperlukan tetapi tidak cukup untuk tindakan yang akan diambil. Masalah terjadi karena skema melibatkan kata-kata seperti 'mungkin', sedangkan logika kondisional bersifat konkret.

Teori skema penalaran pragmatis menjelaskan kinerja penalaran yang buruk dalam hal kurangnya skema yang sesuai untuk tugas tersebut. Tugas abstrak seperti Tugas Inferensi Kondisional terlalu jauh dari pengalaman sehari-hari bagi peserta untuk dapat menanganinya dengan menerapkan skema.

Ini masalah baik untuk teori dan untuk tujuan menemukan mekanisme untuk TFD: pertama, beberapa orang dapat melakukan Tugas Inferensi Kondisional dengan sangat baik, bahkan mereka yang tidak memiliki pelatihan formal dalam logika kondisional dapat mencetak skor di atas tingkat peluang (Inglis & Simpson, 2009a), dan kedua, teori tersebut tidak menyediakan mekanisme potensial yang dapat meningkatkan penalaran abstrak. Masalah terakhir tidak selalu menjadi masalah bagi teori itu sendiri – mungkin tidak mungkin untuk memperbaiki penalaran abstrak – tetapi itu berarti bahwa teori tersebut tidak terlalu berguna dalam konteks saat ini.

#### **1.6.5 Teori proses ganda**

Teori proses ganda mengandaikan bahwa manusia memiliki dua jenis proses kognitif: proses Tipe 1 berbasis heuristik dan proses Tipe 2 berbasis analitik. Proses tipe 1 otomatis, cepat, mudah, tidak disadari, dan lama secara evolusioner. Mereka dibagikan dengan hewan lain dan di antara banyak hal, mereka memungkinkan kita untuk menavigasi lingkungan kita dan menyaring semua informasi yang tidak relevan yang mengelilingi kita (Evans, 2003). Misalnya, mereka memungkinkan kita berjalan melewati ruangan menghindari rintangan tanpa secara sadar memproses setiap detailnya. Mereka kadang-kadang disebut sebagai The Autonomous Set of Systems (TASS, Stanovich, 2004) yang menekankan poin penting bahwa banyak sistem yang berbeda dicakup sebagai tipe 1 berbasis proses.

Proses tipe 2 di sisi lain disengaja, lambat, menuntut memori kerja, sadar dan lebih baru secara evolusioner. Mereka memungkinkan kita untuk menyelesaikan tugas-tugas baru yang kompleks dan evolusioner, seperti, misalnya, menyiapkan pemutar DVD atau bernalar secara hipotetis (Evans, 2003). Di sini, proses akan disebut sebagai Tipe 1 dan Tipe 2.<sup>3</sup>

Teori dual-proses tidak khusus untuk penalaran, melainkan mereka menggambarkan dua jenis pemrosesan yang mendasari semua kognisi. Dengan demikian, berbagai teori proses ganda telah digunakan untuk menyelidiki berbagai fenomena, termasuk penalaran matematis, persuasi sosial, ketakutan akan kematian, memori, penilaian, dan harga diri dan stereotip.

---

<sup>3</sup> Meskipun banyak sumber merujuknya secara berbeda, mis. Sistem 1 dan Sistem 2 (Stanovich, 1999; Evans, 2003), TASS dan analitik (Stanovich, 2004), heuristik dan analitik (Evans, 1984; Evans, 2006), pengalaman dan rasional (Pacini & Epstein, 1999), dan asosiatif dan berbasis aturan (Sloman, 1996).



Disebutkan di atas bahwa beberapa teori proses tunggal dapat dianggap sebagai penjelasan dari salah satu jenis proses dalam teori proses ganda. Teori skema penalaran pragmatis dengan heuristik khusus konteksnya, misalnya, dapat masuk ke dalam proses Tipe 1 di sini, sementara proposisi teori logika mental dari serangkaian langkah logis dapat dilihat sebagai penjelasan tentang bagaimana proses Tipe 2 beroperasi. Hal penting yang perlu diingat di sini adalah berbagai teori tidak harus dilihat sebagai pesaing. Dukungan untuk satu tidak selalu kontradiksi dengan yang lain.

Kembali ke teori dual-proses, masalah penting adalah bagaimana kedua jenis proses berinteraksi. Pada sebagian besar tugas penalaran, kedua jenis proses akan sampai pada kesimpulan yang sama. Ambil contoh silogisme yang valid dengan kesimpulan yang dapat dipercaya:

*Semua ikan bisa berenang.*

*Tuna adalah ikan.*

*Karena tuna adalah ikan maka, tuna bisa berenang.*

Dalam hal ini, jika proses Tipe 1 digunakan, mereka dapat memutuskan bahwa silogisme itu valid karena kesimpulannya dapat dipercaya dan jika proses Tipe 2 digunakan, mereka dapat memutuskan bahwa silogisme itu valid karena secara logis masuk akal. Jadi, output dari sistem akan sama, meskipun pemrosesan yang digunakan untuk menghasilkan output berbeda. Namun, pada beberapa masalah sistem mungkin sampai pada kesimpulan yang berbeda. Silogisme yang valid dengan kesimpulan yang tidak dapat dipercaya pada Gambar 1.8 dapat menunjukkan efek ini.

Sehingga, proses Tipe 1 dapat dengan cepat memutuskan bahwa silogisme itu salah karena kesimpulannya menurut pembaca tidak dapat dipercaya. Proses Tipe 2 yang lambat dan rasional di sisi lain dapat memutuskan bahwa silogisme itu valid karena, seperti di atas, secara logis masuk akal. Teori proses ganda harus menjelaskan bagaimana konflik potensial tersebut ditangani, dan ada beberapa hal yang berbeda dalam penjelasannya: teori persaingan paralel, pre-emptive, dan teori intervensionis standar.

<p><b>Semua hal yang dihisap baik untuk kesehatan. Rokok dihisap. Oleh karena itu, rokok baik untuk kesehatan.</b></p>
--

**Gambar 1.8:** Silogisme yang valid dan tidak dapat dipercaya.

### **1.6.6 Teori persaingan paralel**

Teori persaingan paralel (misalnya Sloman, 1996) menunjukkan bahwa kedua sistem berjalan secara bersamaan dari awal. Jika mereka sampai pada kesimpulan yang sama, pemrosesan berakhir dan kesimpulannya adalah output. Jika prosesnya bertentangan, proses Tipe 2 dapat menimpa proses Tipe 1, tetapi karena proses Tipe 1 jauh lebih cepat, mereka sering memenangkan persaingan. Sloman (1996) menyarankan agar partisipan selalu waspada terhadap kedua respon tersebut ketika ada konflik dan dia menyebutnya sebagai Kriteria S, untuk '*Simultaneous Contradictory Belief*'. Meskipun tanggapan yang bertentangan tidak cukup kuat untuk ditindaklanjuti, keduanya menarik sampai batas tertentu, dan ini menciptakan konflik sadar dalam pikiran si pemberi alasan.

Dari pandangan teori persaingan paralel, ada dua cara di mana penalaran dapat ditingkatkan. Salah satu caranya adalah agar proses Tipe 2 menjadi lebih efisien sehingga lebih sering memenangkan perlombaan dengan proses Tipe 1 dan lebih sering menentukan output. Cara lain adalah untuk heuristik Tipe 1 yang terkadang menyebabkan penalaran yang salah untuk diubah, sehingga pemrosesan Tipe 1 dapat lebih efektif menangani masalah penalaran itu sendiri, dan perlombaan menjadi kurang penting.

Masalah dengan solusi pertama adalah kecepatan sistem secara inheren berbeda sehingga tampaknya tidak mungkin bahwa proses Tipe 2 yang lambat dan membutuhkan sumber daya dapat mencapai titik secepat proses Tipe 1 otomatis dan tanpa usaha (Evans, 2003). Mekanisme yang paling menjanjikan bagi TFD untuk beroperasi melalui teori persaingan paralel adalah perubahan pada beberapa heuristik Tipe 1 yang relevan. Meskipun beberapa heuristik mungkin bawaan, beberapa juga dipelajari dari pengalaman secara bertahap, sehingga masuk akal bahwa studi matematika bisa memiliki beberapa pengaruh bertahap pada mereka.

#### **1.6.7 Teori pre-emptive**

Model pre-emptive berpendapat bahwa tugas diputuskan pada awal, melalui karakteristik superfisialnya, sistem mana yang akan digunakan untuk menyelesaikannya. Ini berarti bahwa tidak pernah ada konflik antara keluaran sistem. Contoh model jenis ini adalah model pengawasan selektif Evans, Newstead dan Byrne (1993), yang dikembangkan untuk memperhitungkan bias keyakinan dalam penalaran silogistik.

Bias kepercayaan terjadi ketika seorang peserta menilai kesimpulan dari silogisme berdasarkan apakah itu dapat dipercaya atau tidak daripada logis atau tidaknya (Evans et al., 1983; S´a et al., 1999). Ambil contoh silogisme yang diberikan di atas (Gambar 1.8) di mana secara logis disimpulkan dari premis bahwa rokok baik untuk kesehatan. Seseorang yang bias oleh keyakinan sebelumnya akan menjawab bahwa silogisme tidak valid, sedangkan seseorang yang mengabaikan keyakinannya dan berkonsentrasi pada langkah logis dari argumen akan menjawab bahwa itu valid. Model pengawasan selektif menunjukkan bahwa 'kepercayaan datang lebih dulu' sehingga ketika sebuah kesimpulan dapat dipercaya, diputuskan di awal bahwa itu benar dan ketika sebuah kesimpulan tidak dapat dipercaya, diputuskan bahwa proses Tipe 2 harus mengevaluasinya lebih teliti.

Sulit untuk mengidentifikasi cara di mana penalaran dapat meningkat di bawah tipe teori pre-emptive. Jika karakteristik tugas menentukan sistem mana yang digunakan, mungkin satu-satunya cara untuk meningkatkan penalaran adalah dengan mengubah heuristik yang mengevaluasi tugas sehingga menjadi lebih konservatif, dan lebih sering menilai bahwa tugas memerlukan pemrosesan Tipe 2.

#### **1.6.8 Teori Default-intervensionis**

Model default intervensionis menyarankan bahwa proses Tipe 1 selalu digunakan sebagai metode default untuk menyelesaikan tugas apa pun, tetapi dalam beberapa kasus, proses Tipe 2 dapat menyimpannya.

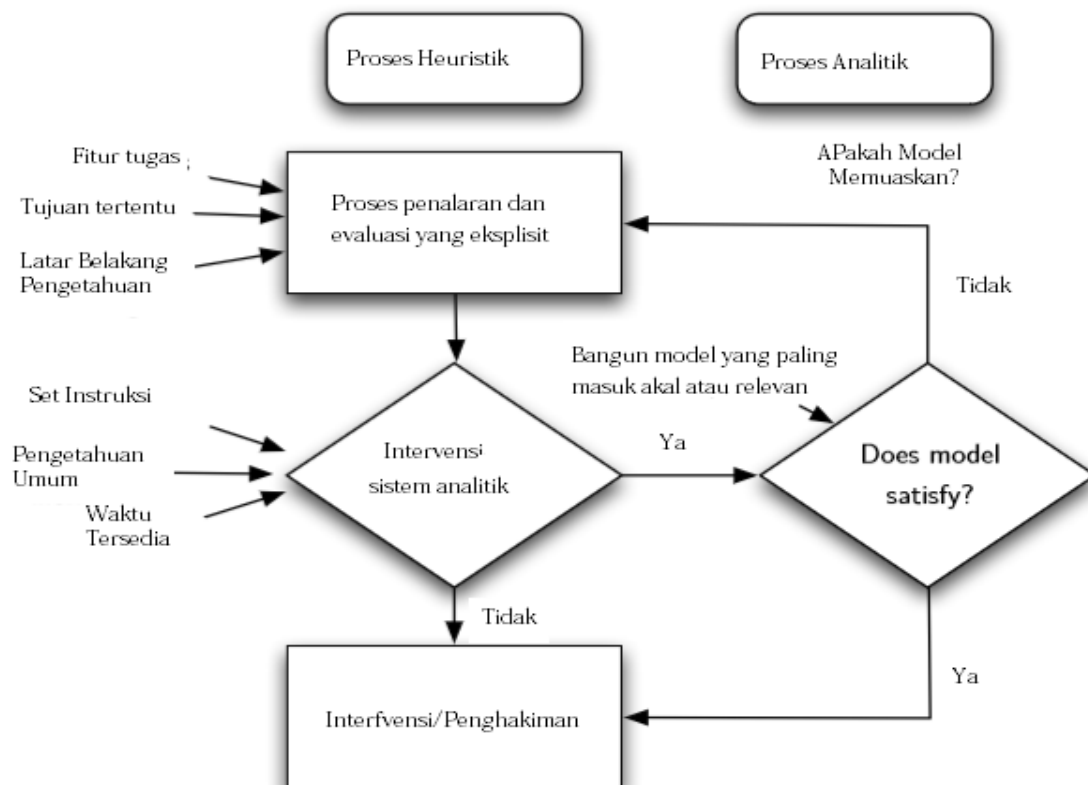
Contohnya adalah teori heuristik-sistematis yang dikemukakan oleh Chen dan Chaiken (1999). Ini mengusulkan bahwa para penalaran memiliki tujuan untuk mengeluarkan upaya minimum yang diperlukan sementara juga berjuang untuk akurasi. Ini berarti mereka menggunakan proses Tipe 1 sebanyak mungkin, tetapi akan menggunakan proses Tipe 2 bila diperlukan untuk kepercayaan diri. Prinsip kecukupan menyatakan bahwa individu memegang kontinum keyakinan tentang penilaian mereka, dengan satu titik pada skala yang berkaitan

dengan keyakinan mereka yang sebenarnya dalam penilaian mereka, dan yang lain berkaitan dengan tingkat kepercayaan yang diinginkan. Penalaran akan menggunakan proses Tipe 1 sejauh mungkin, tetapi jika kepercayaan aktual yang diperoleh dengan ini tidak mencapai tingkat kepercayaan yang diinginkan, mereka akan menggunakan proses Tipe 2 untuk menutup kesenjangan.

Model intervensionis standar lainnya adalah model yang diusulkan oleh Evans (2006). Dalam model Evans (2006) (ditunjukkan pada Gambar 1.9) Tipe 1 memproses keluaran jawaban terlebih dahulu, dan kemudian beberapa bentuk evaluasi terjadi. Seringkali diputuskan bahwa keluarannya masuk akal dan karenanya diterima dan diberikan sebagai tanggapan orang tersebut. Namun, terkadang diputuskan bahwa respons Tipe 1 tidak memuaskan, dan masalahnya diproses ulang dengan proses Tipe 2.

Pendekatan ini mengasumsikan bahwa kita membatasi kognitif sebanyak mungkin, tetapi kita memiliki keinginan untuk merespons dengan benar dan ada mekanisme evaluasi mediasi yang menentukan kapan respons upaya minimum tidak cukup. Perasaan benar (FR, Thompson, 2009) adalah proses meta-kognitif yang dapat menjelaskan bagaimana interaksi antara sistem dimoderasi dalam kasus model default-intervensionis.

Keluaran dari setiap proses kognitif diusulkan untuk datang dengan FR terkait – sebuah intuisi tentang apakah keluaran itu benar atau tidak, mungkin berdasarkan pada kelancaran penghitungan keluaran. Semakin kuat FR, semakin besar kemungkinan bahwa output akan dirasionalisasikan dengan pemrosesan analitik yang dangkal daripada diproses ulang secara analitik. Dalam kasus model default-intervensionis, dapat dianggap bahwa keluaran proses Tipe 1 memiliki FR terkait, dan jika ini tinggi, keluarannya diterima, tetapi jika lebih rendah dari beberapa ambang batas, sistem analitik akan melakukan proses lebih lanjut.



**Gambar 1.9** Model intervensi standar Evans (2006). Model menggunakan istilah 'heuristik' dan 'analitik' untuk merujuk pada proses Tipe 1 dan Tipe 2.

<p>1. Sebuah pemukul dan sebuah bola berharga \$1,10 secara total. Harga kelelawar \$1,00 lebih mahal daripada bola. Berapa harga bolanya?</p> <p>2. Jika dibutuhkan 5 mesin 5 menit untuk membuat 5 widget, berapa lama waktu yang dibutuhkan 100 mesin untuk membuat 100 widget?</p> <p>3. Di sebuah danau ada sepetak bunga lili. Setiap hari ukuran tambalan berlipat ganda. Jika dibutuhkan 48 hari agar petak menutupi seluruh danau, berapa lama petak menutupi setengah danau?</p> <p>Jawaban intuitif: Q1 = 10 sen, Q2 = 100 menit, Q3 = 24 hari. Jawaban yang benar: Q1 = 5 sen, Q2 = 5 menit, Q3 = 47 hari.</p>
--

**Gambar 1.10** Tes Refleksi Kognitif tiga item Frederick (2005).

Dari pandangan model default-intervensionis, ada dua cara di mana matematika dapat meningkatkan kinerja penalaran: mungkin ada beberapa perubahan pada ambang untuk intervensi sistem analitik, sehingga lebih mungkin untuk terlibat; atau mungkin sistem analitik menjadi lebih efisien, sehingga ketika digunakan, kemungkinan besar akan menghitung respons yang benar. Tentu saja, itu juga bisa menjadi kombinasi dari keduanya. Tes Refleksi Kognitif, yang dikembangkan oleh Frederick (2005), memberikan cara yang baik untuk menguji kemungkinan pertama.

Tiga item pada CRT meminta jawaban intuitif, yang dengan cepat muncul di benak tetapi salah. Ini mungkin adalah respons default pemrosesan Tipe 1, dan jika seseorang merespons dengan mereka, diasumsikan bahwa mereka tidak menggunakan proses Tipe 2. Jika seseorang melakukan proses Tipe 2 pada tugas, mereka hampir pasti akan memberikan respons yang benar karena begitu Anda berhenti untuk memikirkannya, mudah untuk melihat bahwa jawaban pertama salah dan mudah untuk menghitung jawaban yang benar (Frederick, 2005). Oleh karena itu, jawaban seseorang untuk setiap pertanyaan memberikan ukuran yang baik apakah mereka telah terlibat atau tidak dalam proses Tipe 2.

Mungkin berguna untuk mendemonstrasikan model intervensionis standar yang ditunjukkan pada Gambar 1.9 menggunakan CRT sebagai contoh. Setelah membaca pertanyaan 1, jawaban '10 sen' muncul di benak - ini adalah model yang paling masuk akal. Karena jawabannya datang dengan FR yang begitu tinggi, kemungkinan akan diputuskan bahwa sistem analitik tidak perlu campur tangan, dan jawaban yang diberikan adalah '10 sen'. Namun, jika seseorang memiliki FR yang lebih rendah, ambang batas yang lebih konservatif untuk FR, lebih cerdas, atau didorong oleh instruksi bahwa tugas tersebut sulit, misalnya, maka mereka dapat memutuskan bahwa sistem analitik harus campur tangan.

Ketika penalaran dan evaluasi eksplisit terjadi, kemungkinan besar kesalahan itu akan terlihat. Seseorang hanya perlu bekerja melalui aritmatika sederhana untuk menemukan masalahnya: 'jika bola berharga 10 sen dan kelelawar berharga \$1,00 lebih banyak daripada bola, maka kelelawar berharga \$1,10, dan secara total harganya \$1,20'. Sekarang masalahnya telah ditemukan dan si pemikir memutuskan bahwa modelnya tidak memuaskan. Model baru (dan biasanya benar dalam kasus CRT, Frederick, 2005) kemudian dibuat, dan jawabannya akan diterima tanpa intervensi analitik lebih lanjut, atau model baru akan dievaluasi secara analitis dan dianggap memuaskan. Jawaban baru (biasanya '5 sen') kemudian diberikan.

Dalam data Frederick (2005) dari 3.428 orang di 11 studi yang berbeda, rata-rata jumlah item CRT yang dijawab dengan benar adalah 1,24 dari 3. Dari semua studi, hanya 17% peserta yang menjawab ketiga pertanyaan dengan benar, 23% menjawab dua dengan benar, 28% hanya menjawab 1 yang benar, dan 33% tidak menjawab pertanyaan dengan benar. Bahkan dalam sampel dengan skor tertinggi, dari Massachusetts Institute of Technology, hanya 48% peserta yang menjawab ketiga item dengan benar. CRT kemudian, sangat sulit. Ini mungkin karena tanggapan intuitif muncul begitu mudah sehingga mereka memiliki FR tinggi dan pertanyaan jarang ditangani oleh proses Tipe 2.

Hal ini dimungkinkan tetapi jarang untuk memberikan respon yang salah non-intuitif (Frederick, 2005), dan dalam hal ini mungkin menunjukkan bahwa proses Tipe 2 telah terlibat, menolak respon default, tetapi gagal untuk menghitung respon yang benar. Meskipun ini tampaknya berguna untuk membedakan keterlibatan sistem analitik dan efisiensi sistem analitik, hal itu terjadi dengan frekuensi yang rendah (karena aritmatikanya sangat sederhana) sehingga tidak mungkin berguna (Frederick, 2005).

Singkatnya, CRT memberikan indikasi yang baik tentang keterlibatan proses Tipe 2. Oleh karena itu, dapat digunakan untuk menguji kemungkinan bahwa belajar matematika meningkatkan kemungkinan intervensi proses Tipe 2. Jika matematika memang meningkatkan kemampuan penalaran tetapi bukan kemungkinan keterlibatan proses Tipe 2, maka mungkin perubahan telah datang dari pemrosesan Tipe 2 yang lebih efisien, menurut perspektif intervensionis default.

### **1.6.9 Membandingkan tiga helai teori proses ganda**

Gillard (2009) melakukan beberapa eksperimen untuk membedakan antara tiga untaian teori proses ganda menggunakan tugas probabilitas. Peserta diperlihatkan gambar dua kotak kelereng hitam dan putih dan ditanya kotak mana yang memberikan probabilitas tertinggi untuk mengambil kelereng hitam. Peserta sering menampilkan apa yang disebut 'pengabaian penyebut' dan memilih kotak dengan frekuensi tertinggi daripada rasio yang lebih disukai.

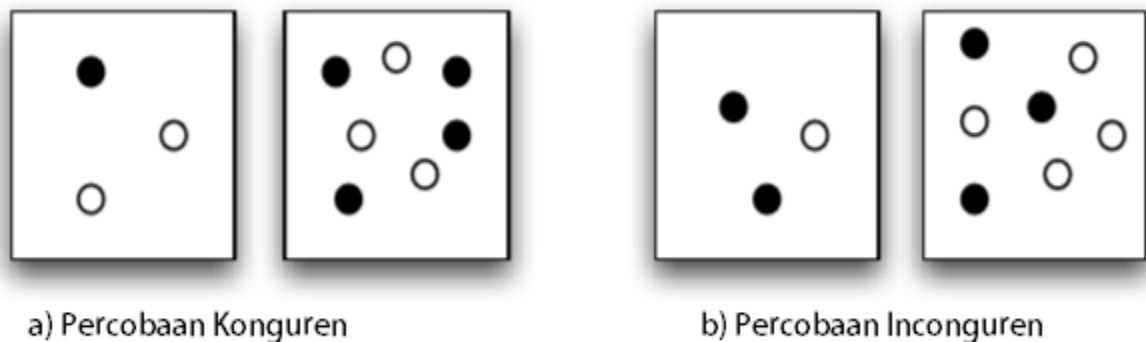
Gillard memvariasikan rasio dan frekuensi absolut untuk membangun percobaan yang kongruen dan inkongruen (Gambar 1.11). Pada percobaan kongruen, kotak dengan peluang terambil hitam tertinggi juga memiliki frekuensi kelereng hitam tertinggi, mis.  $1/3$  dibandingkan dengan  $4/7$ . Pada percobaan yang tidak kongruen, kotak dengan probabilitas tertinggi untuk memilih hitam memiliki frekuensi kelereng hitam paling rendah, mis.  $2/3$  dibandingkan dengan  $3/7$ . Dalam satu percobaan tekanan waktu dimanipulasi, di memori kerja lain dimanipulasi, dan di sepertiga proporsi percobaan kongruen dan inkongruen dimanipulasi, untuk menyelidiki apakah salah satu dari faktor-faktor ini akan mempengaruhi sistem yang digunakan. Waktu reaksi dianalisis dan dibandingkan dengan prediksi yang berasal dari masing-masing teori.

Selama ketiga percobaan, Gillard (2009) menemukan dukungan yang konsisten untuk akun intervensionis default yang dimoderasi FR. Dia berpendapat bahwa tergantung pada kekuatan FR, proses Tipe 2 dapat berkisar dari aktivasi yang sangat minimal dalam bentuk hanya menerima output heuristik, hingga aktivasi yang sangat kuat yang mungkin merupakan perumusan ulang lengkap dari masalah. FR, oleh karena itu, menentukan sejauh mana input sistem analitik, bukan hanya apakah itu akan diaktifkan atau tidak.

Gillard (2009) juga mengusulkan bahwa pemrosesan analitik dapat diperkuat seiring berjalannya waktu. Jika intervensi analitik minimal memunculkan informasi yang

bertentangan dengan model heuristik, pemrosesan analitik akan ditingkatkan. Oleh karena itu, bahkan jika FR awal tinggi dan pemrosesan analitik rendah, masih ada kemungkinan bahwa keluaran heuristik dapat dihilangkan sepenuhnya.

Apa artinya ini untuk tujuan saat ini untuk mengidentifikasi mekanisme yang TFD dapat beroperasi dengan adalah sejauh teori dual-proses penalaran yang bersangkutan, model default-intervensi-dengan-FR mungkin menjadi tempat yang paling menjanjikan untuk dilihat. Sebagaimana dibahas di atas, cara untuk meningkatkan penalaran menurut pandangan ini adalah agar sistem analitik lebih sering digunakan atau agar pemrosesan sistem analitik menjadi lebih efisien saat digunakan. Alasan mengapa sistem analitik digunakan lebih sering adalah karena penurunan umum FR yang terkait dengan keluaran heuristik, atau ambang batas FR yang cukup menjadi lebih konservatif. Salah satu dari perubahan ini akan mengakibatkan sistem analitik lebih sering digunakan. CRT dapat menentukan apakah sistem analitik aktif atau tidak, tetapi tidak harus karena FR rendah atau ambang konservatif. Bahkan mungkin terjadi bahwa tidak mungkin untuk membedakan kedua akun ini secara empiris karena masing-masing akan memiliki efek yang sama pada perilaku yang dapat diamati - dalam kedua kasus diharapkan lebih sering pemrosesan analitik lambat.



**Gambar 1.11:** Contoh a) kongruen dan b) percobaan inkongruen tugas pengabaian penyebut yang digunakan oleh Gillard (2009).

### 1.6.10 Sebuah teori tiga proses

Keith Stanovich telah lama menjadi pendukung pendekatan dual-proses untuk penalaran, tetapi baru-baru ini dia telah melangkah lebih jauh untuk mengusulkan model tripartit (Stanovich, 1999; Stanovich, 2009a). Dia menyarankan bahwa pemrosesan Tipe 2 sebenarnya dapat dibagi menjadi dua tipe lebih lanjut – dalam hal ini, level algoritmik dan level reflektif. Level algoritmik dapat dianggap sebagai elemen komputasi untuk proses Tipe 2, sedangkan level reflektif adalah elemen disposisional (Stanovich, 2009a). Level algoritmik berada di bawah level reflektif di mana level reflektif menentukan kapan level algoritmik akan menimpa proses Tipe 1. Proses tipe 1 dalam teori ini tidak berbeda dengan teori proses ganda yang dibahas di atas.

Teori tripartit dapat dianggap analog dengan teori intervensionis default yang dimoderasi oleh FR yang dibahas di atas. Di sana disarankan agar proses Tipe 1 digunakan secara default dan outputnya memiliki FR terkait. Ketika FR kurang dari beberapa ambang batas, proses Tipe 2 melakukan intervensi untuk melakukan analisis masalah yang lebih teliti. Dalam kasus teori tripartit, proses Tipe 1 lagi-lagi merupakan metode pemrosesan default, tetapi tingkat kognisi reflektif yang disadari daripada FR intuitif yang menentukan kapan harus

menggunakan pemrosesan analitik yang lebih ketat. Ketika level reflektif dianggap perlu, level algoritmik mengambil alih.

Salah satu alasan untuk mengajukan teori tripartit dari struktur ini adalah kemampuan dengan ukuran bernalar kritis telah terbukti agak terpisah dari kecerdasan umum (Stanovich & West, 2008), dan baik disposisi bernalar maupun kecerdasan umum menjelaskan perbedaan unik dalam kemampuan penalaran. Hal ini menunjukkan bahwa kecerdasan umum dan disposisi bernalar memetakan ke berbagai jenis kognisi – tingkat algoritmik dan reflektif, masing-masing. Ini juga berarti bahwa berperilaku secara rasional bergantung pada lebih dari sekadar kapasitas sistem analitik – tetapi juga tergantung pada disposisi untuk berupaya menggunakan kapasitas algoritmik seseorang, sejalan dengan posisi Meliorist pada rasionalitas manusia.

#### **1.6.11 Implikasi untuk TFD**

Dengan asumsi bahwa memang ada tiga jenis pemrosesan dalam penalaran – heuristik, algoritmik, dan reflektif – melalui jenis mana TFD dapat beroperasi? Kemungkinannya adalah (a) mempelajari matematika mengubah heuristik pemrosesan Tipe 1, (b) mempelajari matematika meningkatkan efisiensi algoritme atau menambahkan algoritme baru, dan (c) mempelajari matematika mengubah tingkat reflektif untuk membuat individu lebih tertarik untuk berupaya dalam tugas penalaran.

Jika heuristik Sistem 1 adalah akar dari perbedaan antara perilaku penalaran matematikawan dan non-matematika, maka metode terbaik untuk mengidentifikasi ini adalah pelacakan mata, waktu reaksi, atau tugas yang dipercepat. Dengan cara ini dimungkinkan untuk memisahkan intuisi awal dari pemrosesan analitik yang sedikit lebih lambat, sedangkan tugas berbasis akurasi yang tidak dipercepat seperti tugas Inferensi Kondisional standar dan Silogisme Bias Keyakinan tidak memungkinkan perbedaan seperti itu.

Dengan pelacakan, sangat diharapkan untuk mengidentifikasi aspek dari tugas mana yang menarik perhatian peserta lebih dulu dan mampu menahannya paling lama, dan dengan waktu reaksi dimungkinkan untuk menyimpulkan jenis pemrosesan mana yang menentukan jawaban, berdasarkan kecepatannya. diberikan. Demikian pula, tugas yang dipercepat memungkinkan kita untuk melihat bagaimana peserta akan merespons ketika mereka tidak punya waktu untuk secara efektif terlibat dalam proses Tipe 2.

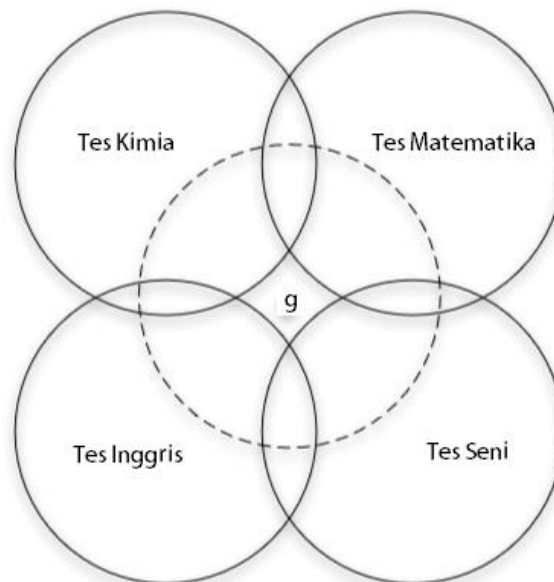
Jika belajar matematika dapat mengubah intuisi seseorang ketika dihadapkan dengan tugas penalaran, tentunya kita akan berharap dapat melihat perbedaan antara ahli matematika dan non-matematika dalam akurasi inferensi kondisional bahkan ketika setiap item memiliki batas waktu yang singkat. Jika matematikawan mengungguli yang lain ketika mereka hanya memiliki, katakanlah, 5 detik untuk menjawab setiap soal, kita dapat menyimpulkan bahwa keunggulan mereka terletak pada pemrosesan Tipe 1.

Jika TFD beroperasi melalui level algoritmik, kita perlu mengidentifikasi ini melalui konstruksi seperti kecerdasan umum dan fungsi eksekutif. Kecerdasan umum adalah aspek mendasar dari kognisi - ini berkontribusi pada kemampuan individu dengan tugas kognitif apa pun. Ketika individu menyelesaikan serangkaian tes yang mengukur kemampuan kognitif yang berbeda, cenderung ada korelasi besar antara kinerja masing-masing, dan ini dianggap karena faktor tunggal yang mendasari kecerdasan umum (*g*, Spearman, 1927; Jensen, 1998, lihat Gambar 1.12).

Kecerdasan dan pendidikan sangat terkait. Mereka yang memiliki kecerdasan lebih tinggi cenderung berprestasi lebih baik di sekolah dan lebih lama bersekolah (Neisser et al.,

1996). Melihat hubungan ke arah lain, kita melihat bahwa sekolah formal mengembangkan kemampuan intelektual dan sekolah di lembaga yang lemah dapat memiliki efek negatif yang signifikan pada skor kecerdasan. Sebagaimana dicatat dalam Bagian 1.3, orang-orang dengan kecerdasan yang lebih tinggi juga mendapatkan derajat yang lebih baik (Farsides & Woodfield, 2003), memiliki kinerja dan kesuksesan pekerjaan yang lebih tinggi, menerima pendapatan yang lebih tinggi (Ashenfelter & Rouse, 1999), dan hidup lebih lama.

Dari sini terlihat bahwa intelegensi dapat menjadi salah satu faktor dalam pengembangan kemampuan penalaran. Sudah diketahui bahwa mahasiswa yang belajar matematika mungkin memiliki kelompok kecerdasan rata-rata yang lebih tinggi daripada mereka yang belajar bahasa dan bahwa skor kecerdasan terkait dengan penalaran kondisional materi (Inglis & Simpson, 2009a). Jika TFD beroperasi melalui level algoritmik Stanovich (2009a), kecerdasan tampaknya merupakan mekanisme yang mungkin – mungkin mempelajari matematika meningkatkan kecerdasan seseorang dan pada gilirannya mengubah perilaku seseorang dengan tugas penalaran seperti *Tugas Inferensi Kondisional*.



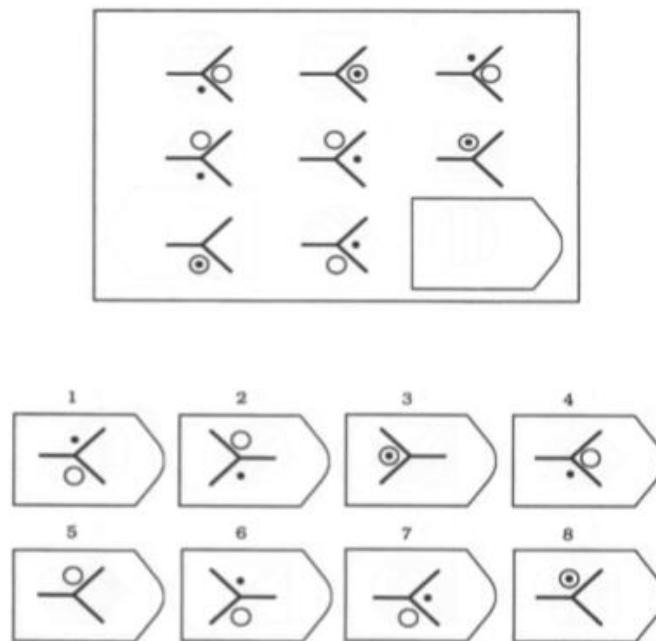
**Gambar 1.12:** Ilustrasi  $g$  sebagai varians bersama antara kemampuan kognitif, seperti mata pelajaran sekolah.

Meskipun beberapa kelenturan dalam kecerdasan terbentuk karena sekolah, umumnya disepakati dalam literatur kecerdasan bahwa  $g$  akan stabil di masa dewasa baik dari waktu ke waktu (Larsen, Hartmann & Nyborg, 2008; Jensen, 1998; Reeve & Lam, 2005; Rönnlund & Nilsson, 2006) dan perubahan lingkungan (Locurto, 1990), jadi meskipun mungkin, kecil kemungkinan belajar matematika pada tingkat lanjut dapat meningkatkan kecerdasan. Namun demikian, jika kecerdasan meningkat melalui studi matematika, cukup masuk akal bahwa ini akan mengarah pada peningkatan keterampilan penalaran umum, sehingga  $g$  adalah faktor yang berharga untuk dipertimbangkan dalam penelitian yang mencari mekanisme peningkatan penalaran.

Matriks Progresif Raven adalah tes kecerdasan non-verbal yang dianggap sebagai ukuran tunggal terbaik dari  $g$  (Jensen, 1998). Tugasnya adalah serangkaian matriks, yang masing-masing menunjukkan pola dengan satu bagian hilang, dan peserta diminta untuk memilih bagian yang hilang dari delapan pilihan pilihan. Contoh item ditunjukkan pada *ketrampilan Penalaran Deduktif (Dr. Agus Wibowo)*



Gambar 1.13. Jika kecerdasan merupakan faktor dalam peningkatan penalaran, maka Matriks Raven harus menjadi ukuran yang baik untuk mencerminkan hal itu. Perlu dicatat bahwa tes kecerdasan, termasuk Matriks Raven, telah terbukti rentan terhadap efek pengujian berulang (Ruston & Jensen, 2010; Bors & Vigneau, 2003). Namun, dalam hal membandingkan peserta dari subjek yang berbeda untuk peningkatan, perbedaan antar kelompok dalam tingkat peningkatanlah yang penting, bukan tingkat peningkatan yang mutlak. Dengan kata lain, pengujian ulang harus mempengaruhi kedua kelompok secara setara (kecuali dalam kasus efek seleksi-maturasi), dan setiap perbedaan antara kelompok dalam tingkat peningkatan dapat dianggap sebagai perubahan nyata dalam gambar 1.13.



**Gambar 1.13:** Contoh item dari Matriks Progresif Tingkat Lanjut Raven. Peserta diminta untuk memilih yang mana dari delapan bagian bernomor yang menyelesaikan grid dengan benar.

Konstruksi tingkat algoritmik lainnya adalah fungsi eksekutif, yang sebenarnya merupakan sekelompok fungsi kognitif: memori kerja (termasuk kapasitas untuk, pemantauan, dan pembaruan informasi yang disimpan dalam pikiran), penghambatan rangsangan yang tidak relevan atau respons dominan, dan pengalihan perhatian di antara tugas-tugas. atau proses (Miyake, Friedman, Emerson, Witzki & Howerter, 2000). Semua kemampuan ini diusulkan untuk terlibat dalam bagaimana kita mengatur proses kognitif kita untuk mencapai suatu tujuan. Ketika kita menghadapi tugas kognitif, khususnya tugas baru, keterampilan ini memandu cara kita mendekati dan menyelesaikannya (Banich, 2009). Ada perselisihan mengenai apakah ketiga fungsi tersebut memiliki mekanisme dasar yang sama (Banich, 2009), dan meskipun Miyake et al. (2000) menemukan korelasi moderat antara kinerja pada setiap aspek, mereka menyimpulkan bahwa mereka sebenarnya dapat dipisahkan dengan jelas.

Stanovich (2009a) berpendapat bahwa fungsi eksekutif itu sendiri bukanlah kata "Eksekutif". "Eksekutif" berarti tingkat pemrosesan tertinggi dan, menurut teori tripartit, tingkat refleksi. Namun, tugas fungsi eksekusi sebenarnya membatasi tingkat refleksi pemrosesan, karena tugas selalu sangat sederhana dan dilengkapi dengan instruksi.

Kriterianya adalah seberapa efektif seseorang dapat mengingat bahwa mereka tahu persis apa yang harus dilakukan.

Tugas tap-perform sebenarnya mengharuskan peserta untuk mencari tahu bagaimana menyelesaikan tugas mereka sendiri dan kemudian menerapkan tingkat usaha yang ditentukan sendiri untuk menyelesaikannya. Seperti disebutkan di atas, tugas fungsi eksekusi mengukur kinerja fungsi tingkat algoritma.

Selain masalah nama dan di samping kecerdasan umum, terdapat tiga fungsi eksekutif utama yang memberikan beberapa wawasan tentang pikiran algoritmik sehingga hal itu merupakan tempat yang berguna untuk mencari perbedaan antara perilaku penalaran matematikawan dan nonmatematika. Moutier, Angeard dan Houd'e (2002), misalnya, menemukan beberapa bukti bahwa pencocokan bias mungkin karena kegagalan penghambatan daripada masalah dengan logika. Ini menunjukkan bagaimana fungsi eksekutif memang bisa menjadi jalur penyelidikan yang menjanjikan ketika mencari mekanisme penalaran yang lebih baik. Jika belajar matematika meningkatkan fungsi algoritmik seseorang, kita mungkin menemukan bahwa tercermin dalam satu atau lebih dari: kecerdasan umum yang lebih tinggi, kapasitas memori kerja yang lebih tinggi, penghambatan yang lebih kuat, atau kemampuan pergeseran yang lebih efisien.

Akhirnya, mungkin TFD beroperasi melalui tingkat reflektif, yang disadap melalui ukuran disposisi bernalar. Stanovich (1999) menggambarkan disposisi bernalar sebagai "mekanisme dan strategi psikologis yang relatif stabil yang cenderung menghasilkan kecenderungan dan taktik perilaku yang khas". Mereka juga dapat disebut sebagai gaya intelektual, gaya kognitif, atau kebiasaan pikiran. Seperti yang telah dibahas sebelumnya, CRT menyediakan ukuran perilaku dari kecenderungan untuk bernalar dengan giat. Langkah-langkah lain yang menilai disposisi bernalar termasuk skala Need for Cognition (NFC) (Cacioppo, Petty & Kao, 1984) dan kuesioner disposisi bernalar yang dirancang oleh Stanovich dan West (1998) yang mencakup subskala bernalar terbuka secara aktif, penalaran kontrafaktual, subskala, subskala absolutisme (diadaptasi dari Erwin, 1981), subskala dogmatisme (diadaptasi dari Troidhal & Powell, 1965 dan Rokeach, 1960), dan terakhir subskala keyakinan paranormal.

Disposisi bernalar seseorang dapat mempengaruhi berapa lama mereka bertahan pada tugas yang sulit, apakah mereka mencari atau menghindari penalaran yang bermanfaat atau seberapa terbuka atau tertutupnya penalaran mereka. Seperti yang diusulkan oleh teori tripartit Stanovich (2009a), disposisi bernalar mungkin sama pentingnya dengan kemampuan kognitif, atau kecerdasan, karena menentukan sejauh mana kecerdasan seseorang akan benar-benar digunakan.

Skala NFC memiliki 18 item laporan diri (Gambar 1.14) yang memberikan ukuran kenikmatan bernalar yang berhasil. Ini jauh lebih pendek daripada kuesioner disposisi bernalar Stanovich dan West (1998), tetapi tetap terkait dengan nilai rata-rata mahasiswa (Elias & Loomis, 2002) dan perilaku pemecahan masalah yang kompleks (Nair & Ramnarayan, 2000).

CRT adalah perilaku yang bertentangan dengan ukuran laporan diri, dan kedua ukuran yang digunakan berdampingan dapat memberikan refleksi yang lebih luas dari disposisi bernalar, menjadikannya ukuran yang cocok untuk studi keterampilan penalaran pada mahasiswa. Jika kasusnya belajar matematika meningkatkan keterampilan penalaran dengan mempengaruhi pikiran reflektif individu, maka kita mungkin menemukan bahwa Need for Cognition atau skor CRT mereka juga meningkat.

### 1.6.12 Ringkasan bagian

- Teori psikologi penalaran cenderung menggambarkan satu, dua atau tiga jenis proses.
- Karya Gillard (2009) mendukung model penalaran default-intervensi. Menurut model penalaran tripartit, yang termasuk dalam kategori default-intervensionis, ada beberapa mekanisme yang memungkinkan TFD beroperasi.
- TFD dapat beroperasi melalui proses Tipe 1, mungkin meskipun mengubah heuristik yang terkadang menyebabkan kesalahan dalam penalaran.
- Ini juga dapat beroperasi melalui tingkat algoritmik dari proses Tipe 2, yaitu mempelajari aturan logis baru, peningkatan kecerdasan, atau peningkatan fungsi eksekutif yang menghasilkan pemrosesan yang lebih efektif atau lebih efisien.
- Akhirnya, TFD dapat beroperasi melalui tingkat reflektif dari proses Tipe 2, yaitu perubahan dalam disposisi bernalar yang berarti individu lebih bersedia untuk terlibat dalam penalaran yang berusaha dan karenanya menggunakan pemrosesan tingkat algoritmik lebih sering atau lebih berhasil.
- Kemungkinan-kemungkinan ini tidak saling eksklusif. Jika TFD benar, ia dapat beroperasi melalui kombinasi mekanisme ini.

## 1.7 STATUS TEORI DISIPLIN FORMAL SAAT INI

Terlepas dari banyak perubahan dalam pendapat psikologis dan bukti yang saling bertentangan tentang apakah keterampilan bernalar dapat digeneralisasikan di seluruh konteks atau tidak, pandangan TFD masih dipegang oleh banyak akademisi dan pembuat kebijakan saat ini. Oakley (1949) berpendapat bahwa “Studi matematika tidak dapat digantikan oleh aktivitas lain apa pun yang akan melatih dan mengembangkan kemampuan logika murni manusia ke tingkat rasionalitas yang sama” dan Amitsur secara serupa menyatakan bahwa “Melalui matematika kita juga berharap untuk mengajarkan penalaran logis – sejauh ini belum ada alat yang lebih baik untuk itu”

1. Saya lebih suka masalah yang rumit daripada yang sederhana.
2. Saya suka memiliki tanggung jawab menangani situasi yang membutuhkan banyak pemikiran.
3. Berpikir bukanlah ide saya untuk bersenang-senang.\*
4. Saya lebih suka melakukan sesuatu yang membutuhkan sedikit pemikiran daripada sesuatu yang pasti menantang kemampuan berpikir saya.\*
5. Saya mencoba mengantisipasi dan menghindari situasi di mana ada kemungkinan saya harus berpikir secara mendalam tentang sesuatu.\*
6. Saya menemukan kepuasan dalam berunding dengan keras dan berjam-jam.
7. Saya hanya berpikir sekeras yang saya harus.\*
8. Saya lebih suka memikirkan proyek kecil harian daripada proyek jangka panjang.\*
9. Saya menyukai tugas yang membutuhkan sedikit pemikiran setelah saya mempelajarinya.\*
10. Gagasan mengandalkan pemikiran untuk membuat jalan saya ke puncak menarik bagi saya.
11. Saya sangat menikmati tugas yang melibatkan menemukan solusi baru untuk masalah.
12. Mempelajari cara berpikir baru tidak terlalu menggairahkan saya.\*
13. Saya lebih suka hidup saya diisi dengan teka-teki yang harus saya pecahkan.
14. Gagasan berpikir abstrak menarik bagi saya.
15. Saya lebih suka tugas yang intelektual, sulit, dan penting daripada tugas yang agak penting tetapi tidak membutuhkan banyak pemikiran.
16. Saya merasa lega daripada puas setelah menyelesaikan tugas yang membutuhkan banyak usaha mental.\*
17. Bagi saya, sesuatu sudah cukup untuk menyelesaikan pekerjaan; Saya tidak peduli bagaimana atau mengapa itu berhasil.\*
18. Saya biasanya berakhir dengan berunding tentang masalah bahkan ketika itu tidak mempengaruhi saya secara pribadi.

**Gambar 1.14** Skala Kebutuhan Kognisi.

TFD juga telah memotivasi argumen agar matematika menerima status khusus dalam Kurikulum Nasional Indonesia. Dalam laporannya tentang pendidikan matematika menengah di Indonesia, Profesor Adrian Smith menyatakan bahwa selain penting untuk kepentingannya sendiri, matematika juga penting karena “mendisiplinkan pikiran, mengembangkan penalaran logis dan kritis, dan mengembangkan analitis dan pemecahan masalah. keterampilan ke tingkat yang tinggi” (Smith, 2004, halaman 11). Berdasarkan hal ini, laporan tersebut merekomendasikan agar prioritas lebih tinggi diberikan untuk "mendorong dan mendanai peningkatan yang signifikan dalam jumlah lulusan matematika yang diterima di Skema Jalur Cepat" (Smith, 2004, hlm. 45), sebuah program yang memberi guru baru titik tulang belakang tambahan pada skala gaji dan beamahasiswa Rp. 5.000.000.

Hal ini adalah salah satu dari banyak rekomendasi yang dimaksudkan untuk meningkatkan status matematika dan meningkatkan jumlah mahasiswa dan guru matematika. Baru-baru ini, laporan Walport (2010) mengidentifikasi penalaran kuantitatif dan logis sebagai keterampilan ilmu, teknologi, teknik dan matematika (STEM), dan mengutip Komite Penasihat Pendidikan Matematika yang menyatakan bahwa "kemampuan pemecahan masalah, ketekunan dan logika adalah [...] sangat dicari dan umumnya ditemukan pada mereka yang memiliki kompetensi tingkat tinggi dalam matematika" (hal.185).

Argumentasi yang dikutip di atas adalah bukti bahwa keyakinan TFD masih dipegang, baik dalam tulisan akademis maupun dalam perdebatan kebijakan. Bahkan telah ditemukan bahwa orang dengan tingkat A dalam matematika memperoleh 7-10% lebih banyak daripada mereka yang memiliki kemampuan dan kualifikasi yang sama pada usia 33 tahun (Dolton & Vignoles, 2002). Kualifikasi matematika jelas sangat dihargai dan kemungkinan hal ini disebabkan, setidaknya sebagian, karena keyakinan bahwa belajar matematika meningkatkan kemampuan penalaran seseorang.

TFD tidak didukung oleh semua. Buku terbaru 'Mengapa Belajar Matematika?' (Bramall & White, 2000) menantang TFD dalam beberapa bab. Bramall secara eksplisit berpendapat bahwa matematika tidak layak untuk diprioritaskan dalam kurikulum dan mata pelajaran harus lebih seimbang (Bramall, 2000). Peter Huckstep meninjau beberapa teori tentang bagaimana matematika melatih pikiran dan menyimpulkan bahwa versi yang paling dapat dipertahankan adalah versi Thomas Tate, yang hanya berlaku untuk tingkat dasar – bukan untuk tingkat lanjutan yang menjadi fokus perdebatan kebijakan baru-baru ini (Huckstep, 2000). Tate menyarankan bahwa matematika memberikan pengenalan yang berguna untuk penalaran tetapi hanya melatih mahasiswa dalam bagaimana menangani pengetahuan matematika - dia tidak percaya bahwa matematika memberikan keterampilan penalaran untuk konteks lain (Huckstep, 2000).

Terlepas dari kenyataan bahwa perdebatan TFD masih belum terselesaikan, kebijakan pendidikan Indonesia dipengaruhi dengan cara yang mengasumsikan TFD benar. Paling-paling, kurangnya bukti konklusif berarti bahwa TFD mungkin tidak memiliki pengaruh sebanyak mungkin, dan paling buruk, itu membuka kemungkinan bahwa itu tidak benar. Oleh karena itu, tujuan dari penelitian ini adalah untuk memberikan tes langsung TFD dengan menyelidiki apakah, dan jika demikian, bagaimana mempelajari matematika di tingkat A dan tingkat sarjana meningkatkan keterampilan penalaran.

### **1.7.1 Ringkasan bagian**

- Kepercayaan pada Teori Disiplin Formal telah dipegang sejak zaman Plato dan masih dipegang sampai sekarang, meskipun ada beberapa perselisihan dari para psikolog akhir-akhir ini.
- Teori Disiplin Formal digunakan dalam laporan kebijakan modern untuk mendukung prioritas matematika dalam kurikulum.
- Penelitian ini bertujuan untuk memberikan tes langsung TFD yang mungkin dapat menginformasikan perdebatan tentang tempat matematika dalam kurikulum nasional Indonesia.

## **BAB 2**

### **DASAR METODE PENELITIAN**

#### **2.1 PENDAHULUAN**

Tujuan dari bab ini adalah untuk menguraikan pertimbangan yang diperlukan untuk melakukan penelitian yang berkualitas tinggi. Area yang dicakup adalah etika penelitian, desain eksperimental dan kuasi-eksperimental, reliabilitas, dan validitas. Potensi masalah di masing-masing bidang ini akan diidentifikasi dan ditangani untuk meletakkan dasar yang kuat bagi penelitian yang disajikan nanti dalam penelitian.

#### **2.2 GAMBARAN UMUM STUDI LONGITUDINAL**

Untuk menempatkan beberapa pertimbangan metodologis yang dibahas di bawah ke dalam konteks, akan sangat membantu untuk memasukkan gambaran singkat dari studi utama yang disajikan kemudian dalam penelitian. Studi utama, yang disajikan dalam Bab 4 dan 5, dirancang untuk menjawab pertanyaan apakah belajar matematika pada tingkat lanjutan dikaitkan dengan peningkatan keterampilan penalaran logis. Studi mengikuti desain kuasi-eksperimental longitudinal. Dalam Bab 4 mahasiswa yang belajar matematika tingkat AS dibandingkan dengan mahasiswa yang belajar Bahasa tingkat AS untuk peningkatan keterampilan penalaran logis selama tahun studi mereka dan di Bab 5, mahasiswa matematika dan psikologi dibandingkan untuk pengembangan keterampilan penalaran selama setahun. Jenis desain ini mengangkat beberapa masalah, yang dibahas pada bagian di bawah ini.

#### **2.3 ETIKA PENELITIAN**

Hal pertama yang mungkin ada di benak ketika subjek etika penelitian muncul sebagai kebutuhan untuk memperhatikan kesejahteraan peserta, kita tidak dapat menundukkan mereka pada kondisi apa pun demi menjawab pertanyaan penelitian kita dan ini dibahas di bawah ini. Pertama, masalah hubungan antara ilmu dan masyarakat dipertimbangkan (Diener & Crandall, 1978) karena relevansinya dengan penelitian tentang TFD.

#### **2.4 ILMU PENGETAHUAN DAN MASYARAKAT**

Masalah hubungan antara ilmu pengetahuan dan masyarakat menyangkut sejauh mana masyarakat harus mempengaruhi topik penelitian yang dikejar (Christensen, 2000). Persaingan untuk pendanaan penelitian bisa sangat ketat, dan uang cenderung digunakan untuk proyek-proyek yang 'sedang mode'. Misalnya, masyarakat kita saat ini prihatin dengan kemiskinan di negara-negara dunia ketiga, iklim ekonomi negara kita sendiri, obesitas, dan imigrasi (antara lain). Sejalan dengan itu, *Economic and Social Research Council* (ESRC), salah satu lembaga pendanaan penelitian utama, menyatakan dalam prioritasnya untuk periode 2009-2014 'mengurangi kemiskinan di antara negara-negara termiskin di dunia', 'pemahaman tanggapan individu dan rumah tangga terhadap perubahan iklim ekonomi yang cepat', 'pengurangan obesitas', dan pemahaman 'dinamika migrasi ke dan dari Indonesia'

Artinya, para peneliti dapat mengubah fokus penelitian mereka demi mendapatkan dana. Meskipun tidak selalu salah bagi peneliti untuk menyelidiki apa yang penting bagi masyarakat, terutama ketika dana penelitian berasal dari uang pembayar pajak, hal itu

membahayakan objektivitas ilmu pengetahuan. Secara tradisional, ilmu seharusnya mengungkap dan menjelaskan sifat dunia kita oleh para ilmuwan objektif, tetapi ini mungkin dikompromikan ketika para peneliti harus bersaing untuk mendapatkan dana dengan menyesuaikan pertanyaan penelitian mereka dengan motif sosial atau politik. Seperti yang ditunjukkan oleh Christensen (2000), minat peneliti juga dapat ditentukan oleh pengalaman pribadi. Misalnya, seorang peneliti disleksia mungkin ingin menyelidiki disleksia pada anak-anak sekolah, dan sekali lagi, ini dapat membahayakan objektivitas penelitian.

Isu masyarakat yang mempengaruhi ilmu pengetahuan relevan dengan penelitian yang dilaporkan dalam penelitian ini. Sebagaimana dibahas dalam Bagian 2.7 tentang status TFD saat ini, ada banyak argumen baru-baru ini agar matematika diprioritaskan dalam Kurikulum Nasional berdasarkan asumsi bahwa mempelajari matematika meningkatkan keterampilan bernalar. Meneliti kebenaran asumsi tersebut merupakan bagian dari motivasi dilakukannya penelitian ini. Oleh karena itu, penting bahwa penelitian dilakukan seobjektif mungkin – tidak membiarkan hasil baik mendukung atau menyangkal TFD. Salah satu cara di mana masalah ini ditangani adalah dengan menahan tujuan penelitian dari para peserta (dibahas di bawah). Manfaat lain dari objektivitas adalah saya mendekati pertanyaan dengan latar belakang psikologi, bukan matematika, dan karena itu tidak merasa perlu untuk mendukung matematika dengan menemukan TFD benar, sedangkan orang lain dengan latar belakang matematika mungkin.

## **2.5 KESEJAHTERAAN PARTISIPAN**

Partisipan dalam studi penelitian harus selalu dimintai persetujuan, diberikan hak untuk mengundurkan diri, dan tanya jawab setelah partisipasi. Jika memungkinkan, penelitian harus dilakukan dengan cara di mana peserta tidak mengalami efek negatif, seperti kegagalan, stres atau rasa malu, dan mereka tidak boleh tertipu. Dua dari masalah ini yang memiliki kepentingan khusus untuk penelitian yang disajikan dalam penelitian ini adalah persetujuan dan penipuan.

### **2.5.1 Izin**

*Informed consent* berarti bahwa peserta harus diberitahu semua informasi yang relevan tentang sifat dan tujuan penelitian dan bahwa mereka hanya boleh mengambil bagian setelah memberikan persetujuan sukarela mereka. Ketika peserta adalah orang dewasa atau anak-anak yang rentan, pertimbangan khusus diperlukan untuk persetujuan yang diinformasikan. Para peserta dalam semua studi yang dilaporkan dalam penelitian ini, termasuk dalam studi tingkat AS, berusia 16 tahun atau lebih. Menurut BPS (The British Psychological Society, 2010), hanya mereka yang berusia di bawah 16 tahun yang dianggap sebagai anak-anak. Namun demikian, persetujuan orang tua diperoleh untuk peserta berusia di bawah 18 tahun di bab 5. Para peserta dalam semua penelitian yang dilaporkan di sini diberi informasi tentang tujuan dan sifat penelitian yang diminta untuk mereka ikuti (kecuali untuk beberapa pemotongan kecil dari informasi dalam kasus studi tingkat AS, lihat di bawah), dan semua memberikan persetujuan sebelum partisipasi mereka dimulai.

### **2.5.2 Penipuan dan penggelapan informasi**

Dalam studi longitudinal yang dilaporkan dalam penelitian ini, perlu untuk menahan beberapa rincian tujuan studi dari peserta. Jika mereka sadar bahwa saya sedang menguji hipotesis bahwa mahasiswa studi matematika akan meningkatkan penalaran ke tingkat

yang lebih besar daripada mahasiswa studi bahasa atau psikologi, hasilnya mungkin telah dibatalkan oleh ancaman stereotip.

Ancaman stereotype terjadi dalam situasi ujian ketika seseorang berpotensi memperkuat stereotip negatif tentang kelompok sosial yang mereka ikuti. Misalnya, ketika pria dan wanita diberi tes matematika yang sama dan diberi tahu bahwa tes tersebut menunjukkan atau tidak cenderung menunjukkan perbedaan gender dalam skor, Spencer, Steele dan Quinn (1999) menemukan bahwa perbedaan gender dalam kinerja mengikutinya:

Untuk kelompok peserta yang diberitahu tes tidak menghasilkan perbedaan gender, tidak ada perbedaan gender yang ditemukan, tetapi untuk kelompok yang diberitahu bahwa itu menghasilkan perbedaan gender, skor wanita secara signifikan lebih rendah daripada pria. Dalam studi lain, wanita Asia ditemukan tampil lebih baik pada tes matematika ketika identitas etnis mereka prima, dan lebih buruk ketika gender mereka prima, daripada kelompok kontrol yang tidak memiliki identitas prima (Shih, Pittinsky & Ambady, 1999). Kinerja matematika pria kulit putih juga ditemukan menderita ketika mereka diingatkan bahwa orang Asia cenderung mengungguli orang kulit putih pada tes matematika, dibandingkan dengan kelompok kontrol yang tidak diingatkan stereotip (Aronson, Lustina, Good & Keough, 1999). Contoh-contoh ini menunjukkan betapa kuatnya efek ancaman stereotip – satu pernyataan dari peneliti dapat menciptakan perbedaan yang signifikan dalam kinerja kelompok peserta.

Dalam studi longitudinal yang dilaporkan di sini, ancaman stereotip bisa menjadi masalah jika peserta diberitahu bahwa saya sedang menguji klaim bahwa mahasiswa matematika akan meningkatkan penalaran ke tingkat yang lebih besar daripada mahasiswa bahasa/psikologi - itu bisa cukup untuk membuat perbedaan antara kelompok yang saya uji. Sebaliknya, peserta diberitahu bahwa penelitian ini melihat peningkatan penalaran di atas A-level tanpa menekankan perbandingan antara mata pelajaran. Para peserta tidak tertipu, oleh karena itu, informasi dirahasiakan. Dalam hal ini, biaya yang ditanggung peserta minimal, dan temuannya tidak akan valid jika informasinya tidak ditahan. Sesuai dengan kode etik BPS (Komite Etika dari British Psychological Society, 2009), dianggap bahwa pemotongan informasi ini tepat, dan penelitian ini disetujui oleh Sub-Komite Etika Persetujuan (Peserta Manusia) Universitas Stekom.

## 2.6 METODE EKSPERIMEN

Sebagian besar studi yang dilaporkan dalam buku ini mengikuti desain eksperimental dan kuasi-eksperimental, dan bagian ini dan selanjutnya akan menguraikan masing-masing metode ini.

Metode eksperimen digunakan untuk menguji hubungan kausal yang dihipotesiskan dengan memanipulasi satu atau lebih variabel secara sistematis dan mengukur pengaruhnya terhadap variabel lain. Hal ini dapat terjadi dalam pengaturan laboratorium yang sangat terkontrol atau dalam pengaturan lapangan yang kurang terkontrol. Variabel yang dimanipulasi disebut sebagai variabel bebas, dan yang diukur adalah variabel terikat. Dalam sebuah eksperimen, variabel bebas dibagi menjadi dua kondisi atau lebih, yang dapat berupa kondisi obat dan kondisi plasebo, atau kondisi waktu tampilan rangsangan 250 md, 500 md, dan 750 md, misalnya. Eksperimen kemudian mengukur dan membandingkan variabel dependen dalam setiap kondisi, mis. laporan pasien tentang gejala mereka setelah perawatan.



**Tabel 2.1** Demonstrasi penyeimbang lengkap dan Latin Square untuk satu set tiga tugas. A = diberikan pertama, B = diberikan kedua, dan C = diberikan ketiga.

Pesanan	Komplit								
	1	2	3	4	5	6	1	2	3
Tugas 1	A	A	B	B	C	C	A	C	B
Tugas 2	B	C	A	C	A	B	B	A	C
Tugas 3	C	B	C	A	B	A	C	B	A

Salah satu cara di mana eksperimen dapat berbeda adalah apakah kondisi tersebut diberikan di antara atau di dalam partisipan. Dalam desain antar partisipan, partisipan secara acak ditempatkan pada kondisi yang berbeda. Penugasan acak sangat penting karena itu berarti bahwa setiap perbedaan antar kelompok, selain manipulasi eksperimental, lebih disebabkan oleh variasi acak daripada variasi sistematis. Setiap perbedaan antara kelompok yang ditemukan dalam ukuran dependen karena itu dapat dikatakan karena manipulasi.

Dalam desain dalam-peserta, semua peserta mengalami semua kondisi, dan urutan pelaksanaannya harus seimbang. Penyeimbangan lengkap berarti bahwa semua urutan tugas yang mungkin diberikan (dengan peserta secara acak ditugaskan ke satu urutan masing-masing). Penyeimbang persegi Latin berarti bahwa tugas-tugas disajikan dalam satu set tetapi urutan berputar. Gambar 2.1 menunjukkan penyeimbang persegi lengkap dan Latin untuk satu set tiga tugas. Alasan penyeimbang adalah untuk mencegah efek urutan. Misalnya, jika peserta menjadi lelah menjelang akhir studi, penyeimbangan memastikan bahwa ini tidak mempengaruhi kinerja pada satu tugas saja tetapi menyeimbangkan efek antara semua tugas.

Christensen (2000) menyatakan bahwa ada tiga syarat yang harus dipenuhi untuk desain penelitian eksperimen yang baik. Kriteria pertama adalah desain harus memungkinkan pertanyaan penelitian dijawab. Kriteria kedua adalah variabel asing (variabel yang tidak menarik tetapi mempengaruhi variabel terikat) dikendalikan (juga dikenal sebagai validitas internal. Yang ketiga adalah temuan tersebut dapat digeneralisasikan.

Kriteria pertama, bahwa desain harus memungkinkan pertanyaan penelitian dijawab, tampaknya begitu mendasar sehingga tidak perlu dinyatakan. Namun, bukan tidak mungkin seorang peneliti sampai sejauh ini mencoba menginterpretasikan data sebelum menyadari bahwa kriteria tersebut belum terpenuhi. Misalnya, ambil pertanyaan penelitian 'Apakah intervensi X efektif dalam membantu anak-anak disleksia meningkatkan kecepatan membaca mereka?'.<sup>1</sup>

Pendekatan yang salah untuk menjawab pertanyaan ini adalah dengan mengambil sampel anak disleksia dan mengukur kecepatan membaca mereka sebelum dan sesudah pemberian intervensi X. Misalkan temuan menunjukkan bahwa kecepatan membaca anak-anak lebih cepat setelah intervensi. Apakah ini menjawab pertanyaan penelitian? Tidak, itu hanya menunjukkan bahwa kecepatan membaca anak-anak menjadi lebih cepat dari waktu ke waktu, apakah itu ada hubungannya dengan intervensi tidak mungkin dikatakan tanpa adanya kelompok kontrol. Ingat studi oleh Lehmann (1963) yang menunjukkan bahwa keterampilan bernalar kritis mahasiswa sarjana meningkat selama pendidikan universitas mereka. Disebutkan dalam Bab 1 dasar teori penalaran bahwa ini bisa menjadi perubahan yang terjadi pada semua orang yang kuliah. Karena kurangnya kelompok kontrol, tidak masuk akal untuk berasumsi bahwa perubahan itu ada hubungannya dengan pengalaman pendidikan peserta. Untuk menjawab pertanyaan seperti itu secara memadai, desain membutuhkan

kelompok eksperimen yang menerima intervensi dan kelompok kontrol yang tidak, dan para peserta harus secara acak ditugaskan ke satu kelompok atau yang lain untuk memungkinkan setiap perbedaan yang disebabkan oleh manipulasi.

Kriteria kedua, variabel asing dikendalikan, diperlukan untuk dapat menghilangkan hipopenelitian saingan. Variabel asing adalah sesuatu selain variabel bebas yang mempengaruhi variabel terikat. Tidak baik jika Anda menyimpulkan bahwa peserta yang membaca cerita bahagia mengingat lebih detail daripada mereka yang membaca cerita sedih jika peserta dalam kelompok cerita bahagia lebih cerdas, misalnya. Cara terbaik untuk mengontrol variabel asing adalah dengan memasukkan kelompok kontrol dan secara acak menetapkan peserta ke dalam kelompok. Kondisi kontrol harus identik dengan kondisi eksperimental di semua aspek kecuali tidak menerima manipulasi eksperimental. Dengan cara ini, variabel independen diisolasi sebagai perbedaan antara kondisi dan pertanyaan penelitian tentang variabel tersebut dapat dijawab. Dengan menetapkan peserta secara acak ke kondisi, seharusnya tidak ada variabel yang diketahui atau tidak diketahui yang mempengaruhi satu kelompok lebih dari yang lain, kecuali untuk variabel independen.

Kriteria ketiga dan terakhir adalah generalisasi (juga dikenal sebagai validitas eksternal) – sejauh mana hasil dapat diterapkan di luar penelitian itu sendiri. Generalisasi dibatasi dengan memiliki sampel non-representatif atau situasi eksperimental buatan. Misalnya, jika sampel Anda seluruhnya terdiri dari perempuan, temuannya mungkin tidak digeneralisasi untuk laki-laki. Demikian pula, jika Anda mempelajari faktor-faktor yang memengaruhi daya tarik wajah menggunakan foto di lingkungan lab, hasil Anda mungkin tidak berlaku untuk daya tarik wajah seperti yang terlihat di lingkungan publik alami. Kemungkinan sampel partisipan akan selalu dibatasi dalam beberapa dimensi, tetapi yang penting adalah menyadari batas-batas populasi tempat sampel Anda berada dan seberapa jauh hasil Anda dapat digeneralisasikan saat kita menarik kesimpulan. Artifisialitas lingkungan studi, bagaimanapun, adalah masalah yang lebih menarik.

Keluhan yang sering dikemukakan pada penelitian berbasis lab dan eksperimental adalah pengaturan yang dibuat-buat membuat hasil tidak dapat digeneralisasikan ke pengaturan lain. Keluhan ini sering mencerminkan kesalahpahaman mendasar tentang tujuan penelitian. Secara keseluruhan, penelitian semacam itu tidak dilakukan dengan tujuan untuk menggeneralisasi hasil ke situasi lain, tetapi dilakukan dengan tujuan menguji suatu teori. Dalam penelitian psikologis, teori menggambarkan dan menjelaskan perilaku dunia nyata dan peran eksperimen adalah untuk menguji dan menyempurnakan teori-teori tersebut. Dalam menguji suatu teori, seorang peneliti memperoleh hipopenelitian yang dapat diuji dalam situasi yang terkendali. Hasil percobaan hanya digunakan untuk menerima atau menolak hipopenelitian itu, sehingga menginformasikan peneliti tentang keakuratan teori dari mana hipopenelitian itu diturunkan. Artinya tidak relevan apakah eksperimen tersebut menyerupai kehidupan nyata atau tidak, yang penting eksperimen tersebut dikontrol dengan ketat sehingga hipopenelitian diuji secara akurat – tanpa adanya variabel pengganggu.

## **2.7 METODE EKSPERIMEN QUASI**

Ekspерimen semu berbeda dari eksperimen sejati karena variabel independen tidak ditetapkan secara acak. Mungkin itu adalah karakteristik yang sudah ada sebelumnya seperti jenis kelamin, kebangsaan, jumlah saudara kandung atau subjek gelar, dalam hal ini kondisi tidak dapat ditetapkan atau dimanipulasi oleh eksperimen. Ini mungkin terjadi bahwa peserta

tidak dapat secara acak ditugaskan karena alasan etis, misalnya, jika Anda melihat keefektifan pengobatan kecanduan alkohol, mungkin tidak etis untuk menolak pengobatan untuk pecandu alkohol dengan menetapkan mereka ke kondisi kontrol (Christensen, 2000), meskipun tentu saja ini sering terjadi dalam penelitian medis, mungkin karena pengetahuan yang diperoleh dianggap membenarkan penghentian pengobatan sementara.

Desain studi longitudinal, penelitian ini adalah quasi-experimental. Hipopenelitiannya adalah belajar matematika meningkatkan keterampilan penalaran mahasiswa ke tingkat yang lebih besar daripada mempelajari mata pelajaran lain. Mahasiswa yang sudah memilih untuk belajar matematika dibandingkan dengan mahasiswa yang sudah memilih untuk belajar mata pelajaran lain. Akan menjadi tidak etis, dan praktis tidak mungkin, untuk secara acak menugaskan orang untuk mempelajari tingkat atau derajat A yang berbeda dan oleh karena itu desain eksperimen semu adalah satu-satunya cara untuk menguji hipopenelitian.

Masalah dengan penugasan non-acak ke kondisi adalah variabel asing, dan khususnya, variabel asing pengganggu, tidak dikontrol dengan benar. Variabel asing adalah faktor selain variabel bebas yang mempengaruhi variabel terikat. Sebuah variabel asing menjadi perancu ketika juga secara sistematis bervariasi dengan variabel independen. Jika variabel pengganggu tidak dikendalikan oleh penugasan acak ke kondisi, itu bisa menjadi penjelasan alternatif untuk setiap efek yang ditemukan dan ini menciptakan masalah untuk menentukan penyebab.

Cara untuk mengatasi masalah ini dalam desain kuasi-eksperimental adalah mengukur dan mengontrol secara statistik untuk setiap faktor yang diantisipasi sebagai pembaur. Dalam kasus studi longitudinal, kelompok matematika mungkin memiliki kecerdasan rata-rata yang lebih tinggi daripada kelompok non-matematika (Inglis & Simpson, 2009a), dan kecerdasan dapat mempengaruhi kemampuan dan perkembangan penalaran. Dalam hal ini, cara untuk mengatasi masalah tersebut adalah dengan mengukur kecerdasan partisipan dan secara statistik mengontrol pengaruhnya dalam analisis.

Dalam eksperimen yang benar dengan penetapan acak pada kondisi, sebab akibat dapat ditetapkan karena tidak ada variabel lain yang dapat menciptakan efek yang diamati. Dalam eksperimen semu, hanya mungkin untuk menetapkan penyebab yang masuk akal dengan menetapkan semua hipopenelitian alternatif yang tidak masuk akal. Christensen (2000) mencontohkan orang yang meninggal segera setelah ditabrak mobil. Ada kemungkinan bahwa mereka benar-benar meninggal karena serangan jantung, terlepas dari ditabrak mobil, tetapi itu sangat tidak masuk akal sehingga Anda dapat menyimpulkan bahwa tabrakan itu adalah penyebab kematian.

Di luar masalah sebab-akibat, ada beberapa masalah lain yang terkait dengan desain kuasi-eksperimental. Dalam setiap contoh yang diberikan di bawah ini, hasil penelitian tampaknya mendukung hipopenelitian bahwa intervensi atau pengobatan berhasil. Namun, ada penjelasan alternatif dari hasil yang belum dikesampingkan.

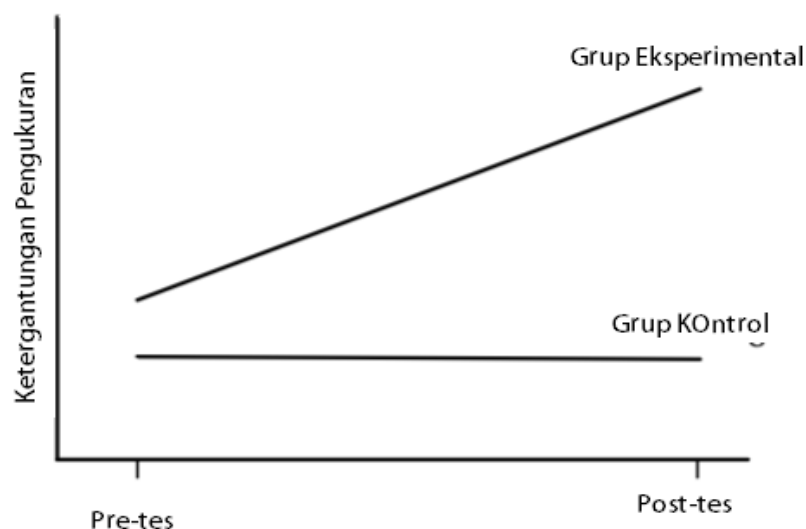
Dalam studi longitudinal yang dilaporkan dalam Bab 5 dan 6 keterampilan penalaran peserta diukur baik sebelum dan sesudah kondisi dialami. Salah satu hasil potensial dari desain ini diidentifikasi oleh Christensen (2000) adalah kelompok pembanding tidak berubah dalam keterampilan penalaran dan kelompok matematika tidak, seperti yang ditunjukkan pada Gambar 2.1. Ini akan menyiratkan bahwa hipopenelitian itu benar, tetapi ada masalah potensial yang dikenal sebagai seleksimaturasi.

Seleksi-maturasi berarti bahwa satu kelompok mungkin sudah berkembang lebih cepat dalam variabel dependen daripada kelompok lain, misalnya, karena mereka lebih cerdas.

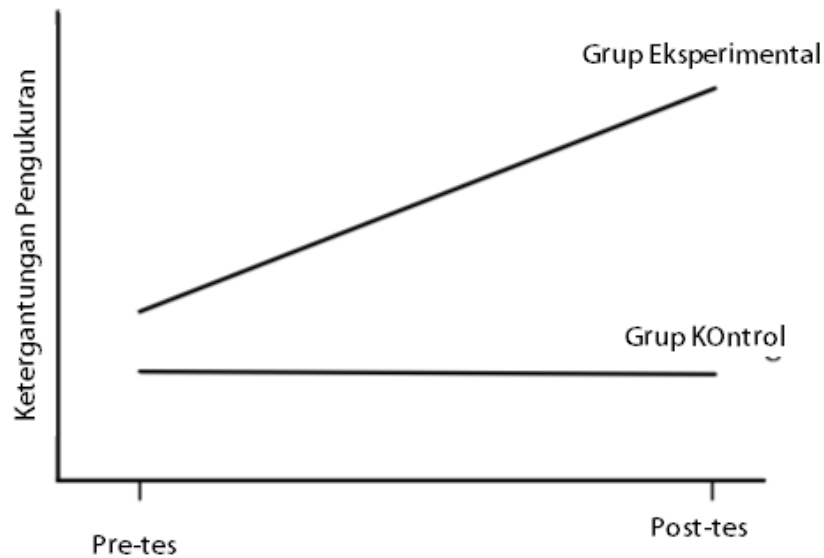
Mungkin mahasiswa matematika lebih cerdas daripada mata pelajaran lain, dan mungkin individu dengan kecerdasan tinggi berada pada lintasan perkembangan yang lebih cepat untuk keterampilan penalaran daripada individu dengan kecerdasan yang lebih rendah.

Ada beberapa pendekatan untuk menangani masalah potensial ini. Salah satunya adalah mencocokkan grup pada variabel asing yang dapat bertanggung jawab atas efek seleksi-maturasi. Dalam hal ini, kita perlu mencocokkan peserta di setiap kelompok pada skor kecerdasan. Ini berarti mengambil peserta dari kisaran yang lebih rendah dari kelompok skor yang lebih tinggi dan dari kisaran atas dari kelompok skor yang lebih rendah sedemikian rupa sehingga kelompok yang dipilih berarti sama dalam ukuran kecerdasan. Ada masalah dengan solusi ini meskipun: peserta yang berada di ekstrem rentang skor kecerdasan kelompok mereka pada pra-tes dapat mundur ke arah rata-rata kelompok mereka dengan pasca-tes, yang dapat membuat kita meremehkan efek dari variabel bebas.

Solusi lain yang mungkin adalah dengan menggunakan metode statistik seperti analisis kovarians (ANCOVA) yang memperhitungkan efek variabel pengganggu saat menentukan hasil. Van Breukelen (2006) membahas keuntungan dan kerugian menggunakan ANCOVA dibandingkan dengan analisis varians tindakan berulang (ANOVA) untuk menyimpulkan efek pengobatan. Metode ANCOVA adalah melakukan analisis terhadap skor Time 2 dengan skor Time 1 sebagai kovariat beserta variabel pengganggu yang dicurigai, sedangkan metode ANOVA pengukuran berulang melibatkan membandingkan perubahan dari awal pada setiap kelompok dengan hanya variabel pengganggu sebagai kovariat. Van Breukelen (2006) berpendapat bahwa dalam studi acak kedua metode tidak bias tetapi ANCOVA memiliki kekuatan lebih. Namun, di mana tidak ada penetapan acak untuk kondisi, ANOVA pengukuran berulang kurang bias karena ANCOVA mengasumsikan tidak ada perbedaan dasar, yang tidak dapat dipastikan dalam desain non-acak. Oleh karena itu, pengukuran berulang ANOVA dengan kovariat kecerdasan dan disposisi bernalar akan digunakan dalam studi longitudinal yang disajikan dalam penelitian ini. Manfaat utama dari pendekatan ini dibandingkan dengan peserta yang cocok adalah sampel tidak berkurang.



**Gambar 2.1** Kemungkinan hasil desain eksperimen semu 1.



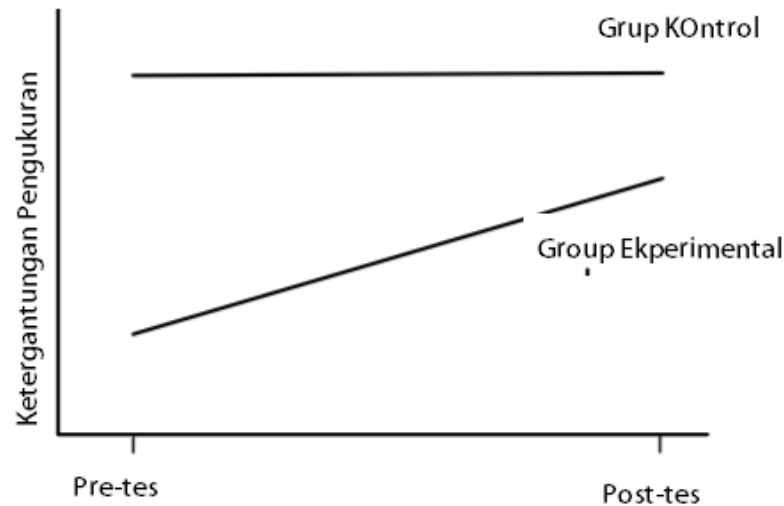
**Gambar 2.2:** Kemungkinan hasil desain eksperimen semu 2.

Selain efek seleksi-maturasi, ancaman potensial lain dengan pola hasil seperti itu adalah efek sejarah lokal. Di sinilah beberapa peristiwa mempengaruhi satu kelompok tetapi tidak yang lain. Beberapa kemungkinan kejadian yang dapat menimbulkan masalah ini dalam studi longitudinal adalah jika kelompok matematika juga lebih mungkin mengambil kursus logika, bernalar kritis, atau mungkin fisika. Ini mudah diidentifikasi dengan mencatat semua mata pelajaran yang dipelajari peserta.

Hasil lain yang mungkin dari desain yang diidentifikasi oleh Christensen (2000), ditunjukkan pada Gambar 2.2, adalah kedua kelompok berubah pada ukuran tergantung dari waktu ke waktu, tetapi kelompok eksperimen berubah lebih dari kelompok kontrol. Sekali lagi, ini bisa menjadi hasil dari efek seleksi-pematangan di mana kelompok eksperimen berada pada lintasan perkembangan yang lebih cepat daripada kelompok kontrol.

Kemungkinan ketiga yang diidentifikasi oleh Christensen (2000) ditunjukkan pada Gambar 2.3. Dalam hal ini skor kelompok eksperimen lebih rendah dari kelompok kontrol pada pre-test, dan meningkat mendekati tingkat kelompok kontrol dengan post-test. Pola hasil ini lebih mungkin terjadi ketika kelompok eksperimen adalah kelompok yang kurang beruntung dan perlakuannya merupakan intervensi yang dirancang untuk membantu mereka. Misalnya, intervensi untuk membantu mahasiswa disleksia meningkatkan kecepatan membaca mereka.

Kemungkinan 3 tidak mungkin terjadi pada kasus kemampuan penalaran pada mahasiswa matematika dan non-matematika, tetapi jika itu terjadi akan ada bahaya bahwa efeknya disebabkan oleh regresi menuju mean dengan nilai yang sangat rendah. Mencetak kelompok eksperimen. Dalam hal ini, perlu juga untuk melacak skor kelompok yang dirampas dari waktu ke waktu tanpa adanya intervensi apa pun. Jika skor konsisten dari waktu ke waktu, itu akan membantu untuk mendukung kesimpulan bahwa peningkatan pada kelompok eksperimen sebenarnya karena pengobatan.



**Gambar 2.3:** Kemungkinan hasil desain kuasi-eksperimental 3.

Kemungkinan keempat dan terakhir yang diidentifikasi oleh Christensen (2000) adalah yang ditunjukkan pada Gambar 2.3 di mana skor kelompok eksperimen mulai di bawah kelompok kontrol dan berakhir lebih tinggi, sedangkan skor kelompok kontrol tidak berubah. Dalam hal ini hipotesis alternatif yang mengancam hasil lain yang mungkin tidak menjadi masalah. Regresi terhadap mean bukanlah kemungkinan yang masuk akal, dan juga bukan efek seleksi-maturasi karena biasanya kelompok yang mendapat skor tertinggi pada pre-test yang berkembang paling cepat.

Untuk menyimpulkan, desain eksperimen semu tidak ideal, tetapi karena peserta tidak dapat secara etis atau praktis ditugaskan untuk mempelajari mata pelajaran yang berbeda di tingkat A atau tingkat gelar, itu harus berhasil. Masalah utama adalah dengan tidak secara acak menugaskan peserta ke kondisi, variabel asing pengganggu tidak dikesampingkan. Ini berarti bahwa mungkin ada penjelasan alternatif untuk setiap efek yang ditemukan. Kecuali jika alternatif-alternatif ini dapat dikontrol secara statistik atau dianggap tidak masuk akal, maka tidak mungkin untuk menetapkan sebab akibat dalam desain eksperimen semu. Faktanya, bahkan jika semua variabel pengganggu yang diketahui dikesampingkan, masih belum aman untuk menyimpulkan hubungan sebab akibat karena mungkin ada variabel pengganggu yang tidak diketahui yang memiliki pengaruh. Penugasan acak ke kondisi benar-benar satu-satunya cara untuk menghindari masalah semacam ini. Ini tidak berarti bahwa temuan studi eksperimen semu tidak penting atau berguna, tetapi itu berarti bahwa temuan itu harus ditafsirkan dengan sangat hati-hati agar tidak melebih-lebihkan hubungan yang ditemukan.

## 2.8 REABILITAS DAN VALIDITAS

Bagian ini akan membahas konsep reliabilitas dan validitas dalam penelitian, khususnya dalam kaitannya dengan penelitian saya sendiri. Secara umum, reliabilitas mengacu pada konsistensi pengukuran, dan validitas mengacu pada sejauh mana suatu instrumen mengukur apa yang ingin Anda ukur. Namun, istilah-istilah ini dapat dibagi menjadi banyak sub-jenis, seperti yang akan terlihat di bawah ini.

### 2.8.1 Reabilitas

Dimulai dengan reliabilitas tes-tes ulang, masalahnya adalah apakah konsep yang Anda ukur diukur secara konsisten sepanjang waktu. Jika, misalnya, Anda mengukur tinggi peserta Anda dua kali dan mendapatkan pengukuran yang berbeda setiap kali, maka kecuali mereka menggunakan obat pemacu pertumbuhan, Anda dapat menyimpulkan bahwa cara Anda mengukur tinggi badan tidak terlalu dapat diandalkan. Kecuali jika konstruk yang Anda ukur diteorikan tidak stabil, maka sebaiknya ukuran Anda memberikan hasil yang sama pada kesempatan yang berbeda.

Lalu, bagaimana Anda dapat memastikan bahwa pengukuran Anda memiliki keandalan yang tinggi? Perubahan pengukuran dapat merupakan hasil dari kesalahan sistematis (faktor perubahan situasi yang membiaskan hasil) atau kesalahan acak. Seperti namanya, kesalahan acak terjadi tanpa alasan tertentu sehingga tidak dapat dikendalikan. Cara untuk meningkatkan reliabilitas, oleh karena itu, adalah dengan ketat mengontrol situasi eksperimental, memastikan bahwa itu adalah semirip mungkin untuk semua peserta setiap kali studi dijalankan.

Cara utama di mana keandalan dimaksimalkan dalam studi longitudinal yang dilaporkan dalam penelitian ini adalah dengan memastikan bahwa semua peserta menyelesaikan tugas dalam kondisi gaya ujian yang sama di setiap titik pengujian. Ini termasuk bekerja sendiri, dalam keheningan, dan memiliki jumlah waktu yang sama untuk menyelesaikan tugas. Kontrol ini harus mengurangi kemungkinan gangguan atau tekanan waktu yang mengganggu kinerja. Akan menjadi masalah jika, misalnya, tugas diselesaikan dalam keheningan pada titik pengujian pertama dan di lingkungan yang bising pada titik pengujian kedua – ini dapat menghambat kinerja putaran kedua memberikan kesan menyesatkan bahwa para peserta telah menjadi sistematis. Lebih buruk dalam penalaran atau bisa menyembunyikan perbaikan nyata. Karena itu, setiap perubahan yang ditemukan harus merupakan produk dari perubahan kemampuan yang sebenarnya, bukan kesalahan dalam pengukuran.

Keandalan internal adalah sejauh mana semua item dalam ukuran terkait, yaitu secara konsisten mengukur konstruk yang sama. Jika Anda memiliki ukuran yang terdiri dari beberapa item, itu adalah ide yang baik bahwa semua item mengukur konstruksi yang sama sampai batas tertentu, meskipun juga diinginkan bahwa setiap item membawa sesuatu yang sedikit unik juga. Ini memastikan bahwa ukuran itu berguna, dapat ditafsirkan, dan tidak terlalu panjang. Sebagai contoh, misalkan Anda memiliki 15 item tugas yang seharusnya mengukur sikap terhadap imigrasi, tetapi tiga item benar-benar mengukur sikap terhadap emigrasi. Sikap terhadap kedua hal tersebut mungkin terpisah dan tidak berhubungan sehingga skor pada 12 item keimigrasian mungkin tidak berkorelasi dengan skor pada ketiga item emigrasi dan ini akan memberikan reliabilitas internal tugas yang rendah. Perhatikan bagaimana hal ini berbeda dari reliabilitas tes-tes ulang – tugas imigrasi dapat memberikan hasil yang konsisten jika dilakukan ulang setiap minggu, tetapi item dalam tugas tersebut tidak menghasilkan respons yang konsisten.

Keandalan internal dapat diukur dengan analisis *split-half* atau *alfa Cronbach* setelah tugas diselesaikan oleh sejumlah peserta. Dalam analisis *split-half*, item pada tugas dibagi menjadi dua kelompok, biasanya dengan percobaan bergantian atau di titik tengah. Skor peserta pada dua bagian item kemudian menjadi sasaran analisis korelasi. Dalam alfa Cronbach, korelasi serupa dihitung tetapi rata-rata di setiap kemungkinan kombinasi bagian

tes. Korelasi yang dihasilkan, atau koefisien reliabilitas, dari salah satu metode dianggap baik jika melebihi 0,8. Keandalan di bawah 0,7 akan menunjukkan keandalan internal yang dipertanyakan, tetapi ini hanyalah aturan praktis. *Split-half* dan analisis *alpha Cronbach* dilaporkan untuk langkah-langkah penalaran yang digunakan dalam bab-bab selanjutnya.

### 2.8.2 Validitas

Validitas adalah konsep yang sedikit lebih luas daripada reliabilitas, mencakup validitas eksternal dan validitas internal. Validitas eksternal adalah sejauh mana sebuah studi digeneralisasikan ke orang lain dan situasi lain, dan itu disinggung di Bagian 3.4. Ini mencakup validitas ekologi - apakah penelitian ini sangat mirip dengan situasi kehidupan nyata - dan validitas temporal - apakah temuan akan berlaku di masa lalu dan masa depan serta masa kini. Penalaran adalah proses kognitif, dan psikologi kognitif mengasumsikan bahwa semua manusia dilahirkan dengan sistem pemrosesan kognitif yang sama. Ini berarti bahwa temuan yang terkait dengan pemrosesan kognitif diasumsikan berlaku lintas waktu dan lintas spesies. Yang mengatakan, ini mengasumsikan bahwa manipulasinya sama. Jika orang lain, di lain waktu, mempelajari silabus matematika tingkat A Indonesia saat ini, maka kita bisa berharap hasilnya sama. Ini bukan untuk mengatakan bahwa hasilnya dapat digeneralisasikan kepada mahasiswa yang mempelajari silabus matematika yang berbeda dalam sistem pendidikan non-Indonesia, atau di masa lalu atau masa depan.

Dalam hal validitas ekologis, satu masalah adalah apakah ukuran kemampuan penalaran yang digunakan dalam penelitian saya relevan dengan klaim TFD tentang jenis penalaran yang dihargai oleh pasar kerja. Apakah tugas berbasis pena dan kertas merupakan ukuran yang valid dari jenis keterampilan penalaran yang mungkin ditunjukkan oleh lulusan matematika dalam pekerjaan mereka di masa depan? Di satu sisi, tidak mungkin bahwa tugas-tugas tersebut sangat mirip dengan tugas-tugas yang akan dihadapi dalam kehidupan sehari-hari. Di sisi lain, dapat dikatakan bahwa peningkatan keterampilan penalaran logis secara umum (seperti yang disarankan oleh TFD) akan ditunjukkan pada tugas penalaran logis apa pun karena sifat universal dari pemrosesan kognitif. Karena TFD tidak merinci secara tepat jenis keterampilan penalaran logis yang ditingkatkan dengan mempelajari matematika, ini adalah yang terbaik yang dapat kita harapkan.

Namun, seperti yang dikatakan Mook (1983), validitas ekologis sering disalahpahami dan tidak diperlukan jika penelitian dirancang untuk menguji teori yang bertentangan dengan generalisasi langsung ke dunia nyata. Dalam penelitian psikologis, validitas ekologis sering dikompromikan demi validitas internal (tidak adanya variabel pengganggu). Hal ini memungkinkan hipopenelitian yang diturunkan dari teori diuji secara efektif. Sama sekali tidak mungkin untuk mengontrol semua variabel pengganggu dalam pengaturan dunia nyata (dan juga sulit untuk menemukan pengaturan dunia nyata yang sama dengan setiap kemungkinan pengaturan dunia nyata yang ingin Anda generalisasikan, Mook, 1983). Dengan secara akurat menguji hipopenelitian yang diturunkan dari sebuah teori dalam pengaturan buatan yang terkontrol, adalah mungkin untuk mendukung, menyangkal atau menyempurnakan teori itu, dan menggunakan teori tersebut untuk menjelaskan perilaku dunia nyata. Dalam penelitian ini, saya menguji TFD, dan teori itulah, bukan data saya, yang digeneralisasikan ke dunia nyata. TFD berpendapat bahwa belajar matematika meningkatkan keterampilan penalaran umum. Jika ini benar, maka kinerja pada tugas 'buatan' yang membutuhkan penalaran harus diubah sebagai bagian dari pengembangan payung.



Seperti yang dinyatakan sebelumnya validitas internal pada dasarnya adalah tidak adanya variabel asing pengganggu. Validitas internal yang tinggi berarti bahwa hubungan antara dua faktor dapat dianggap sebagai hubungan antara kedua faktor itu saja (Heiman, 2002). Jika sebuah penelitian memiliki validitas internal yang rendah, mungkin ada variabel asing yang belum dikendalikan yang mempengaruhi hubungan (membuat mereka variabel pengganggu). Cara terbaik untuk memastikan bahwa sebuah penelitian memiliki validitas internal yang tinggi adalah dengan menetapkan peserta secara acak ke dalam kondisi sehingga setiap variabel asing seimbang dan tidak dapat secara sistematis membiaskan hasil. Upaya perlindungan lainnya adalah mengukur variabel apa pun yang Anda ketahui mungkin merupakan faktor pengganggu sehingga Anda dapat mengontrolnya dalam analisis (Christensen, 2000).

Seperti yang telah dinyatakan, desain kuasi-eksperimental dari studi longitudinal berarti bahwa peserta tidak secara acak ditugaskan ke kondisi. Namun, ada kemungkinan bahwa kecerdasan merupakan faktor pembaur dalam penelitian ini – mungkin berbeda antara kondisi dan ini terkait dengan kemampuan penalaran. Untuk mengatasi masalah ini, kecerdasan akan diukur sehingga efeknya dapat dipertanggungjawabkan dalam analisis.

Validitas internal dapat dikompromikan oleh desain kuasi-eksperimental dari studi longitudinal. Namun, terlepas dari kenyataan bahwa validitas eksternal bisa dibilang tidak relevan karena penelitian ini menguji TFD daripada mencoba menggeneralisasi ke pengaturan dunia nyata secara langsung (Mook, 1983), ia memiliki validitas eksternal yang cukup tinggi karena manipulasi (studi matematika) sedang dilakukan 'di lapangan' daripada di laboratorium, dan karena berurusan dengan proses kognitif yang diyakini umum di seluruh manusia (Miller, 2011).

## BAB 3 MENGUKUR PENALARAN

### 3.1 PENDAHULUAN

Psikologi penalaran telah menjadi bidang penelitian yang kuat dan berkembang sejak tahun 1960-an, ketika Peter Wason pertama kali menunjukkan bahwa orang secara sistematis gagal untuk berperilaku logis pada Tugas Seleksinya yang terkenal. Sejak saat itu, sejumlah besar penelitian telah dilakukan dengan Tugas Seleksi Wason dan kumpulan tugas lainnya. Tujuan dari bab ini adalah untuk membenarkan mengapa tugas Inferensi Kondisional dipilih sebagai ukuran utama penalaran logis dalam penelitian ini, dan mengapa tugas Silogisme Bias Keyakinan dipilih sebagai ukuran sekunder. Untuk melakukan ini, tugas yang paling umum digunakan di lapangan dijelaskan dan didiskusikan.

Bab ini dibagi menjadi dua bagian besar: tugas penilaian dan pengambilan keputusan dan tugas penalaran deduktif. Bagian penalaran deduktif dibagi lagi menjadi tiga bagian: tugas penalaran disjungtif, tugas penalaran kondisional, dan tugas silogisme. Setelah masing-masing tugas dijelaskan, saya akan menyajikan argumen mengapa tugas Inferensi Kondisional dan Silogisme Bias Keyakinan dipilih untuk mengukur penalaran dalam studi yang disajikan dalam penelitian ini.

### 3.2 PENGHAKIMAN DAN PENGAMBILAN KEPUTUSAN

Dalam bagian komentar, Kahneman (1991) mengkarakterisasi bidang penilaian dan pengambilan keputusan dengan tiga fitur: peran teori normatif keyakinan dan pilihan rasional, penekanan pada pilihan berisiko, dan fokus kognitif, daripada sosial atau emosional. Sederhananya, bidang penilaian dan pengambilan keputusan bertujuan untuk menjelaskan dasar kognitif penalaran manusia, dan khususnya, penyimpangannya dari rasionalitas. Aspek terakhir ini dikenal sebagai tradisi heuristik dan bias dan dipelopori oleh Kahneman dan Tversky pada awal 1970-an. Area kerja ini telah berusaha untuk memahami rasionalitas manusia (atau irasionalitas) dengan memeriksa bias yang rentan terhadap kita dan basisnya dalam proses heuristik.

### 3.3 TUGAS HEURISTIK DAN BIAS

#### 3.3.1 Soal hukum bilangan besar

Tversky dan Kahneman (1974) memberi peserta mereka masalah berikut:

*Sebuah kota tertentu dilayani oleh dua rumah sakit. Di rumah sakit yang lebih besar, sekitar 45 bayi lahir setiap hari, dan di rumah sakit yang lebih kecil, sekitar 15 bayi lahir setiap hari. Seperti yang Anda ketahui, sekitar 50 persen dari semua bayi yang lahir adalah laki-laki. Namun, persentase pastinya bervariasi dari hari ke hari. Terkadang mungkin lebih tinggi dari 50 persen, terkadang lebih rendah.*

Untuk jangka waktu satu tahun, setiap rumah sakit mencatat hari-hari di mana lebih dari 60 persen bayi yang lahir adalah laki-laki. Menurut Anda, rumah sakit mana yang mencatat lebih banyak hari seperti itu?

- Rumah sakit yang lebih besar
- Rumah sakit yang lebih kecil

- Hampir sama (yaitu, dalam 5% satu sama lain)

Menurut hukum statistik bilangan besar, semakin besar ukuran sampel semakin dekat mewakili populasi. Dengan aturan ini, rumah sakit besar harus memiliki tingkat kelahiran anak laki-laki yang mendekati 50% daripada rumah sakit kecil. Oleh karena itu, rumah sakit kecil akan memiliki lebih banyak hari di mana lebih dari 60% bayi yang lahir adalah laki-laki. Berbeda dengan hukum jumlah besar, sebagian besar peserta (56%) bernalar bahwa rumah sakit 'hampir sama' dalam hal jumlah hari di mana lebih dari 60% bayi adalah anak laki-laki. Jumlah peserta yang sama (masing-masing 22%) memilih rumah sakit yang lebih kecil atau rumah sakit yang lebih besar. Ini dan banyak penelitian lainnya telah menunjukkan bahwa, secara keseluruhan, orang tidak menggunakan hukum bilangan besar ketika mereka harus.

### **3.3.2 Pengabaian tarif dasar**

Tingkat dasar adalah probabilitas suatu peristiwa yang terjadi tanpa adanya informasi lain. Misalnya, dalam sampel 100 orang di mana 99 adalah perempuan, tingkat dasar perempuan dalam sampel adalah 99/100. Pengabaian tarif dasar adalah masalah lain yang diangkat oleh Kahneman dan Tversky (1972). Masalah ini dijelaskan dengan baik dengan masalah taksi Tversky dan Kahneman (1982):

*Sebuah taksi terlibat dalam tabrak lari di malam hari. Di kota, ada dua perusahaan taksi, Green Cab Company dan Blue Cab Company. Dari taksi di kota, 85% berwarna Hijau dan sisanya berwarna Biru.*

*Seorang saksi mengidentifikasi taksi yang melanggar sebagai Blue. Dalam pengujian di bawah kondisi yang mirip dengan yang terjadi pada malam kecelakaan, saksi ini dengan benar mengidentifikasi masing-masing dari dua warna 80% dari waktu, dan salah 20% dari waktu.*

*Berapa probabilitas bahwa taksi yang terlibat dalam kecelakaan itu ternyata berwarna biru?*

Menurut aturan Bayes, mengingat tarif dasar taksi Biru adalah 0,15 dan saksi mengatakan itu biru dengan akurasi 0,8, kemungkinan taksi itu benar-benar Biru adalah 0,41. Sebaliknya, sebagian besar peserta Tversky dan Kahneman (1982) menilai probabilitas sebagai 0,8, yang merupakan keakuratan saksi. Hal ini menunjukkan bahwa peserta tidak memperhitungkan informasi tarif dasar sama sekali.

### **3.3.3 Masalah Linda**

Masalah Linda yang terkenal berasal dari penelitian Tversky dan Kahneman (1983), dan menunjukkan bias terhadap penalaran bahwa gabungan dua faktor bisa lebih mungkin daripada salah satu faktor saja. Masalahnya berjalan seperti ini:

Linda berusia 31 tahun, lajang, blak-blakan dan sangat cerdas. Dia mengambil jurusan filsafat. Sebagai mahasiswa, ia sangat peduli dengan isu-isu diskriminasi dan keadilan sosial, dan juga berpartisipasi dalam demonstrasi antinuklir.

- 1) Linda adalah seorang guru di sekolah dasar.
- 2) Linda bekerja di toko buku dan mengambil kelas Yoga.
- 3) Linda aktif dalam gerakan feminis.
- 4) Linda adalah seorang pekerja sosial psikiatris.
- 5) Linda adalah anggota Liga Pemilih Wanita.
- 6) Linda adalah seorang teller bank.
- 7) Linda adalah seorang penjual asuransi.
- 8) Linda adalah teller bank dan aktif dalam gerakan feminis.

Peserta diminta untuk membuat peringkat delapan pernyataan yang terkait dengan Linda dalam urutan probabilitas mereka. Temuan yang menarik adalah para peserta cenderung untuk memberi peringkat pada konjungsi 'Linda adalah teller bank dan aktif dalam gerakan feminis' sebagai lebih mungkin daripada bagian konstituen 'Linda adalah teller bank', mungkin karena deskripsi tersebut membuat orang bernalar bahwa Linda harus seorang feminis, apakah dia seorang teller bank atau tidak. Akan tetapi, tidak mungkin sebuah konjungsi dari dua karakteristik menjadi lebih mungkin daripada salah satunya saja.

### **3.3.4 Framming Bias**

Bias pembingkai menggambarkan temuan bahwa peserta dapat memberikan tanggapan yang berbeda terhadap dua pertanyaan yang pada dasarnya sama, tetapi dibingkai secara berbeda. Ambil contoh ini dari Tversky dan Kahneman (1981):

1. Anda adalah petugas layanan kesehatan yang membuat rencana untuk menangani penyakit baru yang akan muncul. Diperkirakan akan membunuh 600 orang. Penasihat ilmiah Anda memberi tahu Anda tentang konsekuensi dari dua kemungkinan program pengobatan: Program A pasti akan menyelamatkan 200 nyawa, sedangkan Program B akan memiliki sepertiga ( $\frac{1}{3}$ ) peluang untuk menyelamatkan 600. Program mana yang akan Anda setujui?
2. Rekan Anda memiliki pilihan antara Program C, yang pasti akan mengakibatkan 400 kematian, dan Program D, yang memiliki peluang dua pertiga ( $\frac{2}{3}$ ) bahwa 600 orang akan mati. Mana yang harus dia setujui?

Pada skenario pertama peserta lebih cenderung memilih Program A, sedangkan pada skenario kedua Program D adalah pilihan yang lebih disukai, meskipun tentu saja A dan C setara dan B dan D setara. Ini dianggap mencerminkan kesulitan terhadap risiko ketika dibingkai dalam hal hasil positif, tetapi preferensi untuk risiko ketika dibingkai dalam hal hasil negatif. Namun demikian, pola tanggapan ini menunjukkan penyimpangan dari logika.

### **3.3.5 Kekeliruan biaya hangus**

Kekeliruan biaya hangus mengacu pada kecenderungan untuk membiarkan biaya hangus sebelumnya (biaya masa lalu yang tidak dapat dipulihkan) memengaruhi pengambilan keputusan saat ini. Misalnya, Anda membeli tiket konser yang tidak dapat dikembalikan, tetapi pada hari itu Anda merasa sangat sakit dan tidak ingin pergi. Jika Anda bernalar Anda harus tetap pergi karena Anda telah membayar tiket, maka Anda akan melakukan kekeliruan biaya hangus.

Toplak, West dan Stanovich (2011) memberi peserta mereka masalah film dari Frisch (1993). Pertama, peserta diminta untuk membayangkan bahwa mereka sedang menginap di kamar hotel, dan mereka baru saja membayar \$6,95 untuk menonton film di TV. Kemudian mereka diberitahu bahwa 5 menit di film itu tampaknya sangat buruk dan mereka bosan. Mereka ditanya apakah mereka akan terus menonton film atau beralih ke saluran lain. Kedua, peserta melihat skenario yang sama kecuali bahwa mereka tidak harus membayar untuk film tersebut. Mereka kembali ditanya apakah mereka akan terus menonton film atau beralih ke saluran lain. Jika peserta melaporkan bahwa mereka akan mengganti saluran saat film tersebut gratis tetapi mereka akan terus menonton setelah mereka membayarnya, mereka dianggap telah melakukan kekeliruan biaya hangus. Hal ini terjadi pada 35,8% peserta dalam studi Toplak et al. (2011).

### 3.3.6 Hasil Bias

Bias hasil adalah kecenderungan untuk menilai kualitas keputusan berdasarkan hasil daripada situasi pada saat keputusan itu dibuat. Masalah yang sering digunakan untuk mengukur bias hasil berasal dari Baron dan Hershey (1988) dan telah digunakan dalam banyak penelitian oleh Stanovich. Peserta diberitahu tentang seorang pria berusia 55 tahun yang memiliki kondisi jantung dan yang menjalani operasi dengan tingkat kematian 8%. Operasi berhasil dan peserta menilai kualitas keputusan pada skala tujuh poin. Kemudian peserta diberitahu tentang seorang pasien dengan kondisi pinggul yang menjalani operasi dengan tingkat kematian 2%. Meskipun keputusan untuk operasi menjadi lebih baik secara objektif, pasien meninggal selama operasi. Jika peserta menilai keputusan pertama (dengan hasil positif) lebih baik daripada yang kedua (dengan hasil negatif), mereka telah menunjukkan bias hasil.

### 3.3.7 Kekeliruan penjudi

Kekeliruan Penjudi mengacu pada kesalahpahaman orang tentang peluang. Seringkali, orang salah percaya bahwa apa yang telah terjadi di masa lalu dapat mempengaruhi kemungkinan kejadian di masa depan. Toplak dkk. (2011) memberi peserta mereka dua masalah yang dirancang untuk memanfaatkan kekeliruan penjudi. Masalah pertama berjalan sebagai berikut:

*Saat bermain mesin slot, orang memenangkan sesuatu sekitar 1 dari setiap 10 kali. Julie, bagaimanapun, baru saja menang dalam tiga permainan pertamanya. Berapa peluangnya untuk menang saat dia bermain lagi?*  
 \_\_\_\_ dari \_\_\_\_

Tanggapan yang benar untuk masalah ini adalah 1 dari 10, peluang yang diberikan dalam pertanyaan. Fakta bahwa Julie telah menang tiga kali tidak berpengaruh pada kemungkinan bahwa dia akan menang pada percobaan berikutnya. Toplak dkk. (2011) menemukan hanya 69,4% tanggapan yang benar untuk masalah ini dalam studi mereka dengan mahasiswa sarjana dan pascasarjana. Masalah kedua yang mereka berikan kepada peserta adalah sebagai berikut:

Bayangkan bahwa kita sedang melempar koin yang adil (koin yang memiliki peluang 50/50 untuk muncul kepala atau ekor) dan itu baru saja muncul 5 kali berturut-turut. Untuk lemparan ke-6 menurut Anda:

1. Ekor lebih mungkin muncul daripada kepala.
2. Lebih mungkin kepala akan muncul daripada ekor.
3. Kemungkinan kepala dan ekor sama pada lemparan keenam.

Jawaban yang benar adalah 3, karena sekali lagi peristiwa masa lalu tidak relevan dengan probabilitas masa depan, dan dalam hal ini 92,2% peserta dalam studi Toplak et al. (2011) menjawab dengan benar, menunjukkan bahwa bias ditampilkan secara tidak konsisten di seluruh tugas.

### 3.3.8 Ringkasan

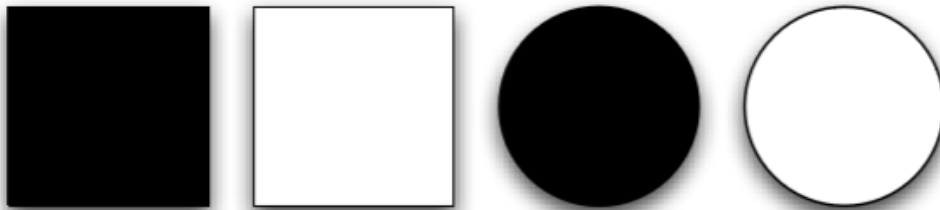
Bagian ini telah menyajikan beberapa tugas umum yang digunakan untuk mengukur bias meresap dalam penilaian manusia dan pengambilan keputusan. Masalah cenderung menyerupai skenario dunia nyata dan masing-masing mengukur aspek kecil dari penalaran manusia (biasanya pada skala biner) yang mungkin penting dalam rentang skenario yang terbatas tetapi tidak selalu relevan secara luas.

Bagian berikutnya membahas penalaran deduktif, yang dapat diukur dengan masalah yang menyerupai dunia nyata, tetapi seringkali tidak. Tugas penalaran deduktif membutuhkan kesimpulan yang diperlukan untuk diturunkan dari premis yang diberikan. Keharusan berarti bahwa kesimpulan harus benar bila premis-premisnya benar. Dengan demikian, penalaran deduktif adalah tentang menilai validitas logis ketika semua informasi yang diperlukan tersedia, bukan tentang membuat keputusan dalam menghadapi informasi yang terbatas.

### 3.4 PENALARAN DEDUKTIF

Penalaran deduktif adalah proses menarik kesimpulan yang diperlukan dari premis yang diberikan. Ini didasarkan pada kepastian mutlak, meskipun premis dapat diasumsikan daripada diketahui. Misalnya, premis 'Semua uskup agung adalah orang percaya' dan 'Tidak ada orang percaya yang kanibal' dapat mengarah pada kesimpulan yang diperlukan bahwa 'Tidak ada uskup agung yang kanibal'. Premis-premis itu mungkin benar atau tidak, tetapi ketika kita menganggapnya benar, kesimpulannya pasti benar. Kepastian inilah yang mendefinisikan suatu deduksi sebagai valid atau tidak valid: suatu deduksi adalah valid jika kesimpulannya harus benar ketika premis-premisnya benar, dan sebaliknya tidak valid. Pada bagian ini dibahas tiga jenis penalaran deduktif: penalaran disjungtif, kondisional, dan silogistik.

Di depan Anda ada empat desain: Kotak Hitam, Kotak Putih, Lingkaran Hitam, Lingkaran Putih



Anda harus berasumsi bahwa saya telah menuliskan salah satu warna (hitam atau putih) dan salah satu bentuknya (persegi atau lingkaran). Sekarang baca aturan berikut dengan seksama.

Jika, dan hanya jika, salah satu desain menyertakan warna yang telah saya tulis, atau bentuk yang telah saya tulis, tetapi tidak keduanya, maka itu disebut THOG.

Saya akan memberitahu Anda bahwa Black Square adalah THOG.

Masing-masing desain sekarang dapat diklasifikasikan ke dalam salah satu kategori berikut:

- A. Pasti adalah THOG
- B. Informasi tidak cukup untuk memutuskan

**Gambar 3.1:** Masalah THOG abstrak Wason.

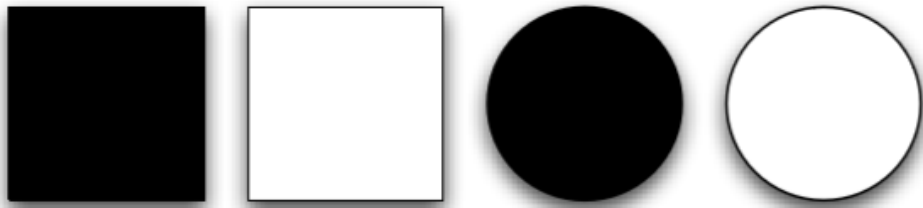
### 3.5 TUGAS DISJUNGSIONAL

#### 3.5.1 Tugas THOG

Masalah disjungsi yang paling terkenal, tugas THOG, dibuat oleh Wason dan ditunjukkan pada Gambar 3.2 (Wason & Brooks, 1979). Pembaca yang tidak terbiasa dengan masalah THOG harus membacanya sekarang sebelum melanjutkan.

Jawaban dari masalah THOG adalah kotak putih dan lingkaran hitam tidak boleh THOG sedangkan lingkaran putih harus THOG, namun hanya 35% partisipan dalam penelitian Wason dan Brooks (1979) yang memberikan jawaban ini. Masalahnya menyatakan bahwa kotak hitam adalah THOG, yang berarti bahwa eksperimen harus memikirkan kotak putih atau lingkaran hitam (THOG berbagi satu karakteristik dengan desain yang dipikirkan eksperimen). Jika pelaku eksperimen memikirkan kotak putih, maka lingkaran hitam (tidak memiliki karakteristik yang sama) dan kotak putih (berbagi keduanya) dapat dikesampingkan sebagai THOG. Jika pelaku eksperimen memikirkan lingkaran hitam maka lingkaran hitam (berbagi keduanya) dan kotak putih (tidak berbagi keduanya) dapat dikesampingkan sebagai THOG. Di bawah kedua alternatif, lingkaran putih memiliki satu karakteristik, dan karenanya merupakan THOG.

Saya telah membawa setumpuk kartu. Ini hanya berisi empat jenis kartu: Kotak Hitam, Kotak Putih, Lingkaran Hitam, Lingkaran Putih



Saya menangani satu untuk diri saya sendiri dari geladak, dan saya tidak akan menunjukkannya kepada Anda. Sekarang saya akan memberi Anda masing-masing kartu, dan saya akan membayar makan malam untuk setiap orang yang memiliki kartu termasuk warna kartu saya, atau bentuk kartu saya, tetapi tidak keduanya.

(Empat kartu di atas diberikan masing-masing kepada Rob, Tim, Paul dan John.)

Tanpa menunjukkan kartu saya, saya dapat memberitahu Anda bahwa saya berutang makan malam kepada Rob. Menurut Anda kartu mana yang bisa saya miliki? Dan apakah Anda pikir saya harus membayar makan malam untuk orang lain? Jika demikian, untuk siapa?

**Gambar 3.2:** Masalah Pub, versi tugas THOG yang dikontekstualisasikan.

Giroto dan Legrenzi (1989) menciptakan masalah pub, sebuah reformulasi dari masalah THOG menggunakan konten yang realistis. Masalahnya adalah tentang karakter bernama Charles yang bermain game dengan empat temannya di sebuah pub. Masalah yang dikemukakan oleh Charles tampak pada Gambar 3.2.

Masalah ini analog dengan masalah abstrak (dalam hal ini jawabannya adalah John juga berutang makan malam), namun 89% orang menjawab dengan benar menurut Giroto

dan Legrenzi (1989). Selanjutnya, ketika Girotto dan Legrenzi (1993) hanya memberi nama SARS pada bentuk yang dihipopenelitiankan dalam versi abstrak tugas, sehingga THOG memiliki satu fitur yang sama dengan SARS, mereka mengamati 70% kinerja yang benar. Penjelasan yang diberikan untuk kesulitan masalah THOG asli disebut teori kebingungan dan berpendapat bahwa orang hanya memperlakukan THOG contoh seolah-olah itu adalah desain yang dipilih oleh eksperimen. Mereka kemudian mencari desain lain yang memiliki satu kesamaan fitur dengan eksemplar tersebut. Disarankan bahwa ketika orang harus mengingat beberapa hipopenelitian sekaligus, seperti dengan disjungsi eksklusif dalam masalah THOG, mereka mengalami kelebihan kognitif dan menggunakan strategi yang lebih intuitif. Dalam hal ini, strategi intuitif adalah mencocokkan nilai-nilai eksemplar dengan kasus uji.

**Tabel 3.1:** Tabel Kebenaran untuk disjungsi eksklusif 'p atau q' di mana T = benar dan F = salah.

$p$	$q$	$p \text{ atau } q$
T	T	F
T	F	T
F	T	T
F	F	F

### 3.5.2 Tugas Tabel Kebenaran

Penalaran disjuntif juga dapat diukur dengan tugas Tabel Kebenaran. Tabel Kebenaran digunakan dalam logika untuk menunjukkan bagaimana kebenaran atau kesalahan setiap variabel menentukan validitas atau ketidakabsahan suatu proposisi tentang variabel-variabel tersebut. Misalnya, Tabel 3.1 menyajikan Tabel Kebenaran untuk aturan disjungsi eksklusif 'p atau q'. Fakta bahwa disjungsi eksklusif berarti p atau q harus benar, tetapi tidak keduanya. Setiap baris mewakili kombinasi kebenaran dan kesalahan yang berbeda dari nilai p dan q, dan kolom terakhir menunjukkan apakah kombinasi tersebut membuat aturan disjuntif benar (valid) atau salah (tidak valid). Dalam disjungsi inklusif, p atau q harus benar, tetapi keduanya juga bisa benar. Tabel Kebenaran untuk disjungsi inklusif ditunjukkan pada Tabel 3.2.

Tugas Tabel Kebenaran yang diberikan kepada peserta untuk mengukur konsepsi mereka tentang kondisi termasuk kebenaran dan kesalahan variabel tetapi biarkan kolom aturan kosong untuk diisi oleh peserta, yaitu peserta memutuskan apakah setiap kombinasi variabel membuat aturan benar atau salah. Ini dapat diberikan dalam bentuk tematik maupun abstrak, seperti yang ditunjukkan pada Tabel 3.3 untuk disjungsi 'Saya akan memesan anggur atau air'. Dengan meminta peserta untuk melengkapi kolom terakhir, kita dapat menyimpulkan seberapa logis orang dalam penilaian mereka tentang disjungsi dan apakah mereka lebih suka interpretasi eksklusif atau inklusif dari disjungsi.

**Tabel 3.2:** Tabel Kebenaran untuk disjungsi inklusif 'p atau q' di mana T = benar dan F = salah.

$p$	$q$	$p \text{ atau } q$
T	T	T
T	F	T
F	T	T
F	F	F



**Tabel 3.3:** Tabel Kebenaran untuk aturan disjungsi 'Saya akan memesan anggur atau air', di mana T = benar dan F = salah.

<i>Aku pesan Es Teh</i>	<i>Aku pesan air putih</i>	<i>Aku akan pesan es teh atau air putih</i>
T	T	T/F?
T	F	T/F?
F	T	T/F?
F	F	T/F?

Evans (1993) meninjau serangkaian studi yang menggunakan tugas Tabel Kebenaran disjungtif abstrak. Dia menemukan bahwa kasus not-p not-q selalu dinilai salah, sebagaimana seharusnya di bawah pembacaan eksklusif dan inklusif, tetapi kasus p not-q dan not-p q dinilai benar sekitar 80% dari waktu, meskipun keduanya benar di bawah kedua bacaan. Hal ini menunjukkan bahwa orang tidak bernalar sepenuhnya secara logis dengan disjungsi. Adapun preferensi untuk bacaan eksklusif atau inklusif, temuannya tidak konsisten. Dalam beberapa studi ada preferensi yang jelas untuk pembacaan eksklusif (di mana kasus pq dinilai salah), di beberapa ada preferensi yang jelas untuk pembacaan inklusif (di mana kasus pq dinilai benar) dan di lain tidak ada preferensi yang jelas.

### 3.5.3 Tugas Inferensi Disjungtif

Dalam tugas Inferensi Disjungtif, peserta diberikan aturan disjungtif bersama dengan premis tentang aturan itu, diikuti oleh kesimpulan yang diturunkan dari aturan dan premis. Partisipan kemudian menilai apakah kesimpulan tersebut valid atau tidak valid. Sebagai contoh:

*Aturan: Premis A atau B*

*Premis: bukan B*

*Kesimpulan: A*

Ada dua kesimpulan penolakan dan dua kesimpulan penegasan. Inferensi penolakan adalah 'Entah p atau q; tidak p; q' dan 'Entah p atau q; bukan q; P'. Kedua kesimpulan ini valid di bawah pembacaan eksklusif dan inklusif. Inferensi afirmasinya adalah 'Entah p atau q; P; bukan q' dan 'Entah p atau q; Q; tidak p'. Di bawah pembacaan eksklusif, kedua kesimpulan afirmasi adalah valid. Di bawah pembacaan inklusif, kesimpulannya mungkin benar atau mungkin tidak, sehingga kesimpulannya tidak valid (belum tentu benar).

**Tabel 3.4** Tabel Kebenaran untuk 'jika p maka q' di mana t = benar dan f = salah.

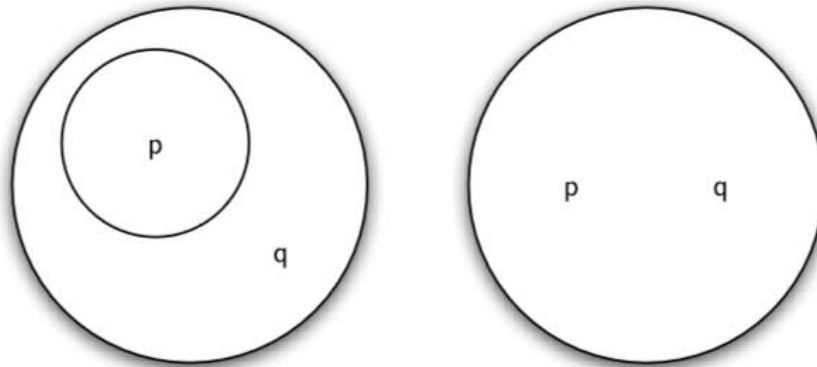
<i>p</i>	<i>q</i>	<i>Materi Kondisional</i>	<i>Bikondisional</i>
t	t	t	t
t	f	f	f
f	t	t	f
f	f	t	t

Seperti tugas Tabel Kebenaran, tugas Inferensi Disjungtif dapat diberikan dengan konten tematik maupun abstrak. Misalnya, 'Kakak saya memelihara ikan tropis, entah itu bidadari atau neon; mereka bukan malaikat; oleh karena itu mereka neon'. Seperti tugas Tabel Kebenaran, jenis tugas ini dapat digunakan untuk menilai sejauh mana peserta menyesuaikan diri dengan logika normatif dan apakah mereka lebih menyukai bacaan eksklusif atau inklusif.

### 3.6 TUGAS KONDISIONAL

Penalaran kondisional adalah proses menarik kesimpulan dari aturan berdasarkan 'jika'. Sebagai Manktelow (1999) berpendapat, kata 'jika' mungkin telah memicu lebih banyak minat dari psikolog, filsuf dan ahli logika daripada kata lain dalam Bahasa.

Aturan kondisional datang dalam berbagai bentuk, termasuk 'p hanya jika q', 'q jika p', 'p jika dan hanya jika q' dan bentuk yang paling umum digunakan, 'jika p maka q'. Menurut logika proposisional formal, masing-masing pernyataan ini harus diperlakukan sebagai kondisional material kecuali untuk 'p jika dan hanya jika q', yang bersifat bikondisional. Di bawah pembacaan kondisional material, aturan hanya terbukti salah jika p benar dan q salah. Di bawah pembacaan bikondisional, p menyiratkan q tetapi q juga menyiratkan p, jadi aturannya salah ketika p salah dan q benar tetapi juga ketika p benar dan q salah. Pembacaan kondisional dan bikondisional material direpresentasikan pada Gambar 3.3 sebagai diagram Euler dan pada Gambar 3.4 sebagai Tabel Kebenaran.



**Gambar 3.3:** Diagram Euler untuk mewakili 'jika p maka q' di bawah a) interpretasi kondisional material, di mana p menyiratkan q, dan b) interpretasi bikondisional, di mana p menyiratkan q dan q menyiratkan p.

#### 3.6.1 Tugas Seleksi Wason

Tugas paling terkenal dalam psikologi penalaran adalah Tugas Seleksi Wason/*Watson Selection Task* (WST), yang dirancang untuk mengukur penalaran kondisional. Hal ini dikembangkan pada 1960-an dan telah melahirkan banyak penelitian sejak, sedemikian rupa, sehingga satu jurnal berhenti menerbitkan penelitian apa pun yang menggunakannya (Manktelow, 1999). Gambar 3.4 menampilkan tugas. Peserta diperlihatkan empat kartu yang masing-masing memiliki huruf di satu sisi dan angka di sisi lain. Dua kartu dengan huruf menghadap ke atas, misalnya A dan D, dan dua kartu menghadap ke atas, katakanlah 3 dan 7. Peserta diminta untuk memilih kartu tersebut, tetapi hanya kartu itu, yang harus mereka balikkan untuk mengetahui apakah aturan seperti 'jika ada A di satu sisi kartu maka ada 3 di sisi lain' adalah benar atau salah. Dalam hal ini, kartu A dan 7 harus dibalik. Jika kartu A memiliki angka bukan-3 di sisi lain, atau jika kartu 7 memiliki A di sisi lain, maka aturannya akan dipalsukan. Namun, ratusan peserta di banyak penelitian gagal membuat pilihan ini. Kebanyakan memilih kartu A dan terkadang kartu 3 juga (Wason & Johnson-Laird, 1972; Evans, 1993).

Bias pencocokan (Evans, 1998) umumnya diamati pada Tugas Seleksi Wason (Wason, 1968). Respons tipikal, A atau A dan 3, masih cenderung diberikan bahkan ketika aturan diubah dari 'jika A maka 3' menjadi 'jika A maka bukan 3', meskipun respons 'A, 3' kemudian menjadi benar secara logis. Hasil ini telah ditafsirkan sebagai indikasi bahwa peserta hanya

mencocokkan kartu dengan aturan daripada menggunakan strategi penalaran sistematis. Wason dan Evans (1975) menunjukkan bahwa peserta tidak menunjukkan kesadaran bias mereka dalam laporan verbal, menunjukkan bahwa itu adalah bias perhatian yang tidak disadari.

WST telah diselidiki dengan konten tematik serta konten abstrak dan dalam banyak kasus, ini telah ditemukan untuk meningkatkan kinerja. Wason dan Shapiro (1971) memberi peserta aturan "setiap kali saya bepergian ke Manchester saya bepergian dengan kereta api", dengan kartu 'Manchester', 'Leeds', 'kereta' dan 'mobil'.

Sementara keberhasilan pada versi abstrak tugas cenderung lebih rendah dari 10%, dalam kasus tematik 10 dari 16 peserta memilih kartu yang benar: Manchester dan mobil. Efek fasilitatif dari konten tematik tidak selalu benar, tetapi tampaknya hanya membantu ketika aturannya bersifat deontik daripada deskriptif, yaitu menyampaikan aturan, izin, atau kewajiban, seperti 'jika seseorang minum alkohol maka mereka harus berusia di atas 18 tahun' atau 'jika seseorang berada di kereta maka mereka harus memiliki tiket yang valid'. Dikatakan bahwa ini karena aturan yang sudah dikenal menghasilkan skema yang dikembangkan secara evolusioner, untuk tujuan seperti pendeteksian penipu (Cosmides, 1989; Cosmides & Tooby, 1992).



**Gambar 3.4:** Contoh Tugas Pemilihan Wason

Terlepas dari popularitasnya, efektivitas WST sebagai ukuran penalaran telah ditantang. Sperber dkk. (1995) telah menyarankan bahwa tugas tidak selalu mengukur penalaran kondisional sama sekali. Sebaliknya, kinerja sangat dipengaruhi oleh mekanisme yang dipandu relevansi yang mendahului mekanisme penalaran apa pun. Ketika dihadapkan dengan masalah penalaran, atau teks lain, pertama-tama kita perlu memahami informasi yang diberikan dan ini termasuk menyimpulkan makna yang dimaksudkan penulis. Dalam kasus tugas inferensi, peserta perlu menyimpulkan atau mengevaluasi kesimpulan yang diturunkan dari premis, jadi meskipun interpretasi mereka terhadap premis mungkin dipengaruhi oleh prinsip relevansi, secara eksplisit jelas bahwa mereka harus melangkah lebih jauh dan terlibat dalam proses penalaran juga. Dalam kasus tugas seleksi, peserta tidak diminta untuk menalar dari premis ke kesimpulan tetapi sebaliknya diminta untuk menilai relevansi masing-masing kartu dengan aturan. Dalam hal ini, penilaian relevansi yang berasal dari proses pemahaman memberikan jawaban intuitif untuk masalah dan tidak ada kebutuhan eksplisit untuk terlibat dalam penalaran lebih lanjut. Ini mungkin menjadi sumber bias pencocokan bias.

**Tabel 3.5:** Tabel Kebenaran untuk 'jika p maka q' dimana T = benar dan F = salah.

<i>p</i>	<i>q</i>	<i>jika p maka q</i>
T	T	T
T	F	F
F	T	T
F	F	T

Interpretasi tugas seleksi ini didukung oleh enam studi di dua makalah (Sperber et al., 1995; Sperber & Girotto, 2002). Sperber dan rekan-rekannya telah menunjukkan bahwa tingkat keberhasilan dalam tugas dapat dimanipulasi secara dramatis dengan mengubah faktor relevansi konten. Keberhasilan dalam versi deskriptif tugas dapat ditingkatkan hingga lebih dari 50%, lebih sesuai dengan tingkat keberhasilan yang biasanya ditemukan dengan versi deontik (Sperber et al., 1995). Selanjutnya, keberhasilan dalam versi deontik dapat dikurangi hingga di bawah 20%, serupa dengan tingkat yang biasanya ditemukan pada versi deskriptif dan abstrak (Girotto, Kimmelmair, Sperber & van der Henst, 2001).

Karena kontroversi ini, WST dikesampingkan sebagai ukuran kemampuan penalaran untuk penelitian ini. Seandainya mahasiswa matematika ditemukan mengalami perubahan dalam kinerja WST di samping studi matematika mereka, tidak akan jelas apakah perubahan itu terjadi pada kemampuan penalaran kondisional atau dalam proses interpretasi.

### 3.6.2 Tabel Kebenaran

Tugas Tabel Kebenaran dibahas dalam bab 3.3.1 sebagai ukuran penalaran disjungtif, tetapi tugas tersebut juga dapat digunakan untuk mengukur penalaran kondisional. Tabel 3.5 menyajikan Tabel Kebenaran untuk aturan kondisional 'jika p maka q'. Sekali lagi, setiap baris mewakili kombinasi kebenaran dan kesalahan yang berbeda dari nilai p dan q, dan kolom terakhir menunjukkan apakah kombinasi itu membuat aturan kondisional benar atau salah. Sekali lagi, tugas Tabel Kebenaran dapat diberikan kepada peserta untuk diselesaikan dalam bentuk tematik maupun abstrak, seperti yang ditunjukkan pada Tabel 3.6 untuk persyaratan 'Jika hujan maka saya bawa payung'.

Dengan meminta peserta untuk memutuskan baris mana dari tabel yang mereka anggap valid, dimungkinkan untuk menentukan pembacaan kondisional mana yang paling cocok dengan interpretasi mereka. Seperti yang dibahas dalam bab 1.2, setidaknya ada empat cara di mana orang dapat menafsirkan pernyataan kondisional: kondisional material (yang didukung oleh logika proposisional formal), bikondisional, kondisional cacat dan kondisi konjungtif, dan ini berbeda dalam hal Tabel Kebenaran (lihat Tabel 1.2).

<p>Jika hurufnya H maka angkanya 5          Hurufnya H          Kesimpulan: Angkanya 5</p> <p><input type="radio"/> Ya</p> <p><input type="radio"/> Tidak</p>
---

**Gambar 3.5:** Contoh item Modus Ponens dari tugas Inferensi Kondisional

### 3.6.3 Tugas Inferensi Kondisional

Dalam tugas Inferensi Kondisional, peserta diberikan aturan kondisional bersama dengan premis tentang aturan itu, diikuti oleh kesimpulan yang berasal dari aturan dan premis.

Partisipan kemudian menyimpulkan apakah kesimpulan tersebut valid atau tidak valid. Sebagai alternatif, peserta dapat menghasilkan kesimpulan yang mereka anggap valid, tetapi ini jauh lebih jarang terjadi dalam literatur dan oleh karena itu versi evaluasi difokuskan di sini.

Gambar 3.5 menunjukkan item tugas inferensi kondisional yang khas dengan kesimpulan yang valid. Ini adalah contoh inferensi Modus Ponens, salah satu dari empat jenis inferensi yang digunakan dalam tugas. Keempat inferensi, *modus ponens* (MP), *denial of the antecedent* (DA), *armation of the consequent* (AC) dan *modus tollens* (MT) ditunjukkan pada Tabel 3.7, bersama dengan empat bentuk aturan yang dibuat dengan memutar kehadiran negasi, dan apakah kesimpulan-kesimpulan tersebut dianggap sah menurut keempat interpretasi tersebut.

Sebuah abstrak (hanya menggunakan huruf dan angka) versi 32-item dari tugas inferensi kondisional digunakan oleh Inglis dan Simpson (2008, 2009a) untuk membandingkan perilaku penalaran mahasiswa matematika dan non-matematika. Tugas tersebut mencakup empat jenis inferensi, masing-masing disajikan empat kali dengan empat bentuk aturan berbeda yang ditunjukkan pada Tabel 3.7, yang dibuat dengan memvariasikan posisi negatif. Ini menciptakan 16 item, dengan negasi eksplisit. Dalam 16 item lebih lanjut, masalahnya identik dalam struktur kecuali bahwa negasinya tersirat (misalnya 'bukan 3' mungkin direpresentasikan sebagai '6'). Gambar 4.6 menunjukkan beberapa contoh item dari tugas.

**Tabel 3.6:** Tabel Kebenaran untuk aturan kondisional 'jika hujan maka saya akan membawa payung', di mana T = benar dan F = salah.

<i>ini hujan</i>	<i>Aku membawa payung</i>	<i>Jika hujan maka aku membawa payung</i>
T	T	
T	F	
F	T	
F	F	

<p>Jika hurufnya S maka angkanya 6            Angkanya bukan 6            Kesimpulan : Hurufnya bukan S</p> <p><input type="radio"/> Ya  <input type="radio"/> Tidak</p> <p>a) Modus tollens</p>	<p>Jika hurufnya M maka angkanya 4            Surat itu bukan M            Kesimpulan: Angkanya bukan 4</p> <p><input type="radio"/> Ya  <input type="radio"/> Tidak</p> <p>b) Penolakan dari pendahulunya</p>
--	--

**Gambar 3.6:** Contoh item dari tugas Inferensi Kondisional yang menunjukkan a) inferensi Modus Tollens dan b) penolakan inferensi Anteseden.

Inglis dan Simpson (2009a) menemukan bahwa mahasiswa matematika mengungguli mahasiswa non-matematika (berdasarkan materi kondisional menjadi bacaan normatif, lihat bab. 2), bahkan ketika kelompok dicocokkan untuk kecerdasan umum. Maahamasiswa matematika tidak meningkat dalam kinerja tugas selama satu tahun, tetapi perbedaan awal

membuka dua kemungkinan: mahasiswa matematika mungkin telah meningkat dalam penalaran kondisional selama pra-universitas tetapi studi pasca-wajib matematika, sejalan dengan TFD, atau bisa jadi orang dengan gaya penalaran yang lebih normatif secara tidak proporsional disaring untuk mempelajari matematika pasca-wajib.

**Tabel 3.7:** Empat inferensi dan jenis pernyataan kondisional dengan dan tanpa premis negasi (Pr) dan kesimpulan (Con). Simbol harus dibaca 'tidak'. Di bagian bawah, validitas setiap kesimpulan di bawah setiap interpretasi diberikan.

Kondisional	MP		DA		AC		MT	
	Pr	Con	Pr	Con	Pr	Con	Pr	Con
jika p maka q	p	q	$\neg p$	$\neg q$	q	p	$\neg q$	$\neg p$
jika p maka $\neg q$	p	$\neg q$	$\neg p$	q	$\neg q$	p	q	p
jika $\neg p$ maka q	$\neg p$	q	p	$\neg q$	q	$\neg p$	$\neg q$	p
jika $\neg p$ maka $\neg q$	$\neg p$	$\neg q$	p	q	$\neg q$	$\neg p$	q	p
<i>Jenis premis minor</i>	<i>Afirmatif</i>		<i>Denial</i>		<i>Afirmatif</i>		<i>Denial</i>	
Validitas Material	Valid		Invalid		Invalid		Valid	
Validitas Defektif	Valid		Invalid		Invalid		Invalid	
Validitas Bikondisional	Valid		Valid		Valid		Valid	
Validitas KonjungSIONal	Valid		Invalid		Valid		Invalid	

Dalam studi wawancara yang dilakukan oleh Inglis (2012), delapan pemangku kepentingan dalam komunitas matematika (misalnya anggota komite pendidikan Institut Matematika dan Aplikasinya dan London Mathematical Society) diminta untuk melihat tugas Inferensi Kondisional dan menilai mereka setuju dengan pernyataan "Tugas ini menangkap beberapa keterampilan yang dikembangkan dalam mempelajari matematika tingkat lanjut". Dari delapan peserta tersebut, enam sangat setuju dengan pernyataan tersebut (lima pada skala Likert lima poin) dan dua setuju (empat pada skala Likert lima poin). Seorang peserta bahkan mengatakan bahwa "Jika belajar matematika tingkat A tidak membuat Anda lebih baik dalam hal itu, ada yang salah dengan silabus".

### 3.7 TUGAS SILOGISME

Silogisme mewakili bentuk tertua dari penalaran deduktif formal, berasal dari zaman Aristoteles lebih dari 2.000 tahun yang lalu (Manktelow, 1999). Mereka dibentuk dari pernyataan empat bentuk: 'Semua A adalah B', 'Beberapa A adalah B', 'Tidak A adalah B', dan 'Beberapa A bukan B', masing-masing dikenal sebagai A, I, E dan O. A dan I afirmatif sedangkan E dan O negatif, dan A dan E universal sedangkan I dan O partikular. Sebuah silogisme seperti 'Semua A adalah B; Semua B adalah C; Semua A adalah C' adalah contoh struktur AAA. Dalam semua silogisme, kesimpulan akan menggambarkan hubungan antara suku pertama dan terakhir dari premis, dalam hal ini, A dan C. Dimungkinkan untuk membangun 512 silogisme yang berbeda, tetapi jumlah yang dianggap valid belum disepakati: angka berkisar luas, antara

14 dan 48 (Manktelow, 1999). Interpretasi manusia terhadap silogisme telah diselidiki secara luas, menggunakan tugas-tugas abstrak dan kontekstual

### 3.7.1 Tugas Silogisme Bias Keyakinan

Tugas yang umum digunakan adalah Tugas Silogisme Bias Keyakinan (Sa et al., 1999). Dalam tugas ini, peserta melihat 24 silogisme, 12 di antaranya valid dan 12 di antaranya tidak valid, dan mereka diminta untuk memutuskan yang mana. Namun, silogisme dibuat menjadi dapat dipercaya, tidak dapat dipercaya atau netral-kepercayaan (ada 8 dari masing-masing jenis, 4 di antaranya valid dan 4 di antaranya tidak valid, contoh masing-masing jenis diberikan pada Gambar 3.7), dan seterusnya. adalah mungkin untuk menentukan sejauh mana partisipan dibujuk oleh keyakinan dan logika (peserta diinstruksikan untuk mengesampingkan keyakinan mereka dan bernalar secara logis, tetapi hal ini sulit dilakukan).

Misalnya, dalam silogisme yang valid dengan konten yang tidak dapat dipercaya, seperti 'Semua hal yang dihisap baik untuk kesehatan; rokok dihisap; oleh karena itu rokok baik untuk kesehatan' seseorang yang dapat mengatasi keyakinan dan alasan mereka sebelumnya dengan logika akan lebih mungkin untuk menerima silogisme sebagai valid daripada seseorang yang lebih banyak terpengaruh oleh keyakinan mereka daripada oleh instruksi untuk mengabaikannya. Apa yang disediakan oleh tugas Belief Bias Silogisme, oleh karena itu, adalah ukuran kemampuan peserta untuk menalar secara independen dari keyakinan mereka sebelumnya. Ini dianggap sebagai komponen sentral dari penalaran kritis.

Dapat dipercaya, Valid :	Dapat dipercaya, Valid :
Premis : Semua ikan dapat berenang Tuna adalah ikan	Premis : Semua yang hidup membutuhkan air mawar membutuhkan air
Kesimpulan : Tuna dapat berenang	Kesimpulan : Mawar adalah benda hidup
Tidak dapat dipercaya, valid :	Tidak dapat dipercaya, valid :
Premis : Semua yang berkaki empat berbahaya Anjing tidak berbahaya	Premis : Semua senapan berbahaya Ular berbahaya
Kesimpulan : Anjing tidak memiliki empat kaki	Kesimpulan : Ular merupakan senapan
Netral, valid :	Netral, valid :
Premis : Semua Biskuit rasanya enak. Oreo adalah ramadion.	Premis : Semua manusia mengenakan pakaian. Andi mengenakan pakaian
Kesimpulan : Oreo rasanya enak.	Kesimpulan : Andi adalah manusia

**Gambar 3.7:** Contoh item dari tugas Belief Bias Silogisme.

Ada dua ukuran yang dapat diambil dari tugas Belief Bias Silogisme: skor total, mencerminkan kemampuan penalaran silogistik lintas jenis item, dan Belief Bias Index (BBI), yang merupakan jumlah total item yang konsisten didukung (valid-believable, invalid-unbelievable) dikurangi jumlah total item yang tidak konsisten yang didukung (valid-unbelievable, invalid-believable) dan mencerminkan kemampuan peserta untuk bernalar berdasarkan validitas logis atas kepercayaan.

Banyak penelitian telah dilakukan dengan tugas Belief Bias Silogisme, dan ada tiga temuan utama: item yang valid diterima lebih sering daripada item yang tidak valid, item yang dapat dipercaya diterima lebih sering daripada item yang tidak dapat dipercaya, dan kepercayaan dan validitas berinteraksi (Manktelow, 1999). Evans dkk. (1983) menemukan bahwa item yang valid diterima lebih sering daripada tidak apakah item tersebut dapat dipercaya atau tidak dapat dipercaya, tetapi item yang tidak valid hanya diterima lebih jarang daripada tidak ketika item tersebut tidak dapat dipercaya; item tidak valid dipercaya salah

diterima 66% dari waktu. Hal ini menunjukkan bahwa ketika silogisme mudah dipercaya, orang cenderung menerimanya tanpa bernalar lebih jauh. Evans dkk. (1983) menyarankan model pengawasan selektif untuk menjelaskan temuan. Idenya adalah orang-orang pada awalnya bernalar dengan heuristik yang memberitahu mereka untuk langsung menerima barang yang dapat dipercaya. Hanya barang-barang yang tidak dapat dipercaya yang diperiksa lebih lanjut dengan pemrosesan Tipe 2 yang disengaja dan disadari.

### 3.8 RINGKASAN

Seharusnya jelas sekarang bahwa ada banyak tugas yang telah digunakan untuk mengukur berbagai aspek penalaran manusia. Tidaklah mungkin untuk menggunakan semua tugas ini dalam studi penelitian dan masih berharap untuk menemukan peserta yang bersedia. Secara khusus, studi longitudinal yang disajikan dalam Bab 4 dan 5 dari penelitian ini dilakukan selama kelas sekolah dan universitas, sehingga ukuran yang digunakan perlu disesuaikan dengan durasi kelas standar (50 menit). Untuk memberikan waktu bagi kovariat untuk diukur serta kemampuan penalaran, diputuskan bahwa hanya dua dari banyak tugas penalaran yang dibahas di atas yang harus dipilih.

Tujuan dari penelitian ini adalah untuk menyelidiki perubahan perilaku penalaran dalam matematika dibandingkan dengan mahasiswa non-matematika, dan untuk menyelidiki kemungkinan mekanisme untuk setiap perubahan yang ditemukan. Telah dikemukakan bahwa penalaran kondisional merupakan pusat penalaran logis dan juga matematika. Houston (2009) berpendapat bahwa sebagian besar pernyataan matematis berbentuk 'jika pernyataan A benar, maka pernyataan B benar', bahkan jika pernyataan tersebut sangat disamakan (hal. 63). Untuk alasan ini, penalaran kondisional dapat menjadi tempat yang berguna untuk memulai penyelidikan keterampilan penalaran pada mahasiswa matematika. Namun, seperti yang terlihat di atas, ada tiga tugas utama yang telah digunakan untuk mengukur keterampilan penalaran kondisional: Tugas Seleksi Wason, tugas Tabel Kebenaran, dan tugas Inferensi Kondisional.

Temuan yang dibahas dalam bab ini menunjukkan bahwa tugas Inferensi Kondisional akan menjadi ukuran yang paling tepat dari penalaran kondisional untuk digunakan dalam penelitian ini. Pertama, tugas tersebut telah digunakan secara luas dan dihormati secara luas dalam literatur psikologi penalaran (bebas dari jenis kritik yang telah diajukan terhadap Tugas Seleksi, misalnya Sperber et al., 1995). Kedua, mengukur aspek penalaran di mana mahasiswa matematika dan nonmatematika telah ditemukan berbeda (Inglis & Simpson, 2008, 2009a). Ketiga, ini akan memungkinkan saya untuk membedakan antara dua hipopenelitian yang dibiarkan terbuka oleh Inglis dan Simpson (2009a): hipopenelitian pengembangan dan hipopenelitian penyaringan. Akhirnya, dianggap oleh sekelompok pemangku kepentingan dalam komunitas pendidikan matematika sebagai ukuran yang baik dari keterampilan belajar matematika yang berkembang (Inglis, 2012).

Selain tugas Conditional Inference, termasuk tugas Belief Bias Silogisme sebagai ukuran penalaran dalam penelitian ini memungkinkan saya untuk memperluas ruang lingkup pekerjaan dalam dua cara: dengan mengukur kemampuan penalaran silogistik serta kemampuan penalaran kondisional, dan dengan mengukur penalaran dengan konten tematik serta penalaran dengan materi abstrak. Tugas Silogisme Bias Keyakinan memberikan ukuran kemampuan penalaran silogistik dan sejauh mana orang dapat bernalar dengan logika atas keyakinan, dan ini merupakan komponen penting dari penalaran kritis.



Ingat kutipan dari laporan Smith (2004) bahwa mempelajari matematika "mengembangkan penalaran logis dan kritis, dan mengembangkan keterampilan analitis dan pemecahan masalah ke tingkat yang tinggi". Tugas Conditional Inference berkaitan dengan aspek logika kutipan ini, sedangkan tugas Belief Bias Silogisme berkaitan dengan aspek bernalar kritis.

## BAB 4

### MENGEMBANGKAN KETRAMPILAN BERNALAR PADA MAHASISWA

#### 4.1 PENGUJIAN TEORI DISIPLIN FORMAL

Tujuan dari penelitian ini adalah untuk menyelidiki teori *Theory of Formal Discipline* (TFD) yang menyatakan bahwa belajar matematika meningkatkan keterampilan penalaran. Dua pertanyaan penelitian utama adalah (a) apakah belajar matematika pada tingkat lanjutan terkait dengan peningkatan keterampilan penalaran dan (b) jika ada peningkatan, bagaimana mekanismenya? Beberapa cahaya akan diberikan pada kedua masalah ini dalam bab ini, meskipun itu tidak berarti bahwa keduanya akan diselesaikan secara meyakinkan.

TFD menunjukkan bahwa belajar matematika meningkatkan penalaran logis dan keterampilan bernalar kritis seseorang. Keyakinan ini dipegang oleh para filsuf (Locke, 1971/1706; Plato, 2003/375B.C), matematikawan (Amitsur, 1998; Oakley, 1949), dan pembuat kebijakan (Smith, 2004; Walport, 2010). Sampai sekarang, bagaimanapun, bukti empiris yang mendukung TFD sangat jarang. Ada bukti campuran apakah keterampilan bernalar dapat ditransfer sama sekali (lihat Bab 2 untuk tinjauan bukti), dan meskipun ada bukti keterampilan penalaran yang lebih baik dari mereka yang telah mempelajari matematika di tingkat lanjutan daripada mereka yang tidak (Inglis & Simpson, 2009a), tidak ada bukti keterampilan penalaran berkembang bersama studi matematika. Studi yang disajikan di sini menyelidiki apakah belajar matematika di tingkat AS (tahun pertama tingkat A) dikaitkan dengan peningkatan yang lebih besar dalam penalaran logis dan keterampilan bernalar kritis daripada mempelajari bahasa. Peningkatan keterampilan penalaran didefinisikan sebagai penalaran yang lebih dekat dengan model normatif yang relevan.

Ukuran penalaran logis dan bernalar kritis yang akan saya gunakan adalah tugas Inferensi Kondisional dan Silogisme Bias Keyakinan. Alasan untuk ini diuraikan dalam Bab 4, tetapi secara singkat, penalaran kondisional dianggap penting bagi penalaran logis (Anderson & Belnap, 1975; Braine, 1978; Inglis & Simpson, 2008), dan kemampuan untuk memisahkan keyakinan sebelumnya dari validitas logis. dianggap sebagai bagian dari bernalar kritis.

Jika ditemukan bahwa belajar matematika memang terkait erat dengan peningkatan yang lebih besar dalam keterampilan penalaran daripada mempelajari bahasa, itu akan menimbulkan pertanyaan bagaimana mekanisme peningkatan tersebut. Model tripartit Stanovich (2009a) diperkenalkan dalam tinjauan literatur sebagai titik awal untuk mengidentifikasi mekanisme potensial untuk peningkatan keterampilan penalaran. Model tripartit adalah perpanjangan dari model penalaran dual-proses yang mengusulkan proses Tipe 1 yang cepat dan otomatis dan proses Tipe 2 yang lambat dan disengaja (Evans, 2003). Dalam model tripartit, pemrosesan Tipe 2 dikatakan terjadi pada dua tingkat – tingkat algoritmik dan reflektif. Level algoritmik adalah elemen komputasi untuk pemrosesan Tipe 2 – kapasitas yang tersedia untuk pemrosesan yang berhasil dan efisiensi yang dengannya pemrosesan yang berhasil dapat terjadi. Level reflektif adalah elemen disposisional – pemrosesan yang mengatur kapan dan sejauh mana level algoritmik akan digunakan sebagai lawan dari pemrosesan Tipe 1.

Ada kemungkinan bahwa peningkatan keterampilan penalaran dapat disebabkan oleh perubahan pada salah satu dari ketiga jenis proses ini. Mempelajari matematika dapat

mengubah proses Tipe 1 yang memusatkan perhatian kita pada aspek-aspek tertentu dari suatu masalah. Atau dapat meningkatkan kapasitas tingkat algoritmik untuk atau efisiensi penalaran yang berhasil. Akhirnya, bisa jadi bahwa belajar matematika mengubah tingkat reflektif yang membuat orang yang bernalar lebih mau berusaha untuk bernalar. Atau untuk faktor domain-umum, sumber peningkatan dalam penalaran bisa menjadi apa yang disebut Stanovich (2009a) "mindware". Mindware terdiri dari pengetahuan, aturan, dan prosedur spesifik domain yang dapat dipanggil kembali secara eksplisit dari memori untuk membantu memecahkan masalah spesifik, daripada berguna untuk penalaran yang lebih umum. Mungkin pengetahuan, aturan atau prosedur diajarkan dalam matematika yang membantu para pemikir ketika memecahkan jenis tugas tertentu.

Pemrosesan tipe 1 paling baik dipelajari dengan waktu reaksi atau ukuran akurasi terbatas waktu karena kecepatan dan otomatisitasnya (Gillard, 2009; Evans & CurtisHolmes, 2005). Metode tersebut memungkinkan tanggapan awal untuk diisolasi dari tanggapan selanjutnya yang didasarkan pada pemrosesan Tipe 2, tetapi mereka memerlukan administrasi berbasis komputer dan dengan demikian tidak cocok untuk penelitian yang dilaporkan dalam bab ini (di mana semua tindakan diberikan di atas kertas di sekolah). Sebaliknya, peran pemrosesan Tipe 1 dalam perbedaan antara perilaku penalaran matematikawan dan non-matematika dieksplorasi dalam Bab 8. Di sini, tingkat algoritmik dan reflektif dari pemrosesan Tipe 2 diselidiki sebagai mekanisme potensial untuk peningkatan keterampilan penalaran.

Tingkat algoritme dapat dinilai melalui ukuran kecerdasan dan fungsi eksekutif, yang mencerminkan kapasitas dan efisiensi kognitif. Hubungan antara fungsi eksekutif dan keterampilan penalaran diselidiki secara terpisah di Bab 8, sementara kecerdasan dipertimbangkan di sini. Jika belajar matematika dikaitkan dengan peningkatan keterampilan penalaran karena peningkatan kapasitas untuk pemrosesan Tipe 2 maka ini mungkin tercermin dalam peningkatan skor pada tes kecerdasan. Subset item dari Matriks Progresif Lanjutan Raven (RAPM, Raven et al., 1998) dimasukkan sebagai ukuran kecerdasan dalam studi utama yang disajikan di bawah ini untuk memungkinkan mekanisme semacam itu diidentifikasi. Sebagaimana dicatat dalam Bab 1, RAPM adalah tugas penyelesaian pola non-verbal yang dianggap sebagai ukuran tunggal terbaik dari kecerdasan umum (Jensen, 1998). Sebuah subset terbatas waktu item yang dipilih untuk populasi mahasiswa telah terbukti memiliki keandalan split-setengah yang dapat diterima dari 0,79 (Stanovich & Cunningham, 1992). Hal ini juga memungkinkan perbedaan antar-kelompok dalam kecerdasan pada Waktu 1 untuk dikendalikan (lihat diskusi tentang metode eksperimen semu di Bab 2).

Tingkat kognisi reflektif dinilai dengan ukuran disposisi bernalar, seperti skala Bernalar Bernalar Terbuka Aktif (AOT, Stanovich & West, 1997), skala Kebutuhan untuk Kognisi (NFC, Cacioppo et al., 1984) dan Kognitif Uji Refleksi (CRT, Frederick, 2005). Tes Refleksi Kognitif, yang diperkenalkan pada Bab 2, mengajukan tiga pertanyaan yang mendorong tanggapan intuitif tetapi salah. Peserta perlu menghambat tanggapan ini untuk menjawab dengan benar. Baru-baru ini, Toplak et al. (2011) menunjukkan bahwa CRT adalah prediktor yang lebih baik dari penalaran normatif daripada kecerdasan, fungsi eksekutif, atau AOT. Hal ini menunjukkan dua hal: pertama, jika belajar matematika dikaitkan dengan penalaran yang lebih normatif, maka perubahan ke tingkat reflektif mungkin merupakan mekanisme yang lebih mungkin daripada perubahan ke tingkat algoritmik, dan kedua, tingkat reflektif mungkin lebih baik dimanfaatkan oleh ukuran kinerja (misalnya CRT) daripada ukuran laporan diri (misalnya

timbangan AOT dan NFC). Di sini, tingkat reflektif akan diukur dengan dua tugas: CRT, dan skala NFC. Skala NFC tidak termasuk dalam penilaian Toplak et al. (2011) dan dapat memberikan kekuatan prediksi tambahan.

Sampai saat ini, belum ada saran tentang sifat dari mindware matematis apa pun yang bertanggung jawab untuk perbaikan, jadi titik awalnya adalah dengan hanya melihat efek subjek yang dipelajari (matematika atau non-matematika) dan apakah ini memprediksi peningkatan secara independen dari faktor domain-umum. Jika faktor umum domain tidak ditemukan sebagai prediktor, maka mindware matematika adalah kemungkinan yang tersisa.

Sebagai poin tambahan, jika matematika mengembangkan keterampilan penalaran melalui mekanisme domain-umum maka kita mungkin mengharapkan peningkatan yang luas dalam kinerja penalaran. Jika, di sisi lain, matematika menyediakan beberapa mindware khusus, kita mungkin mengharapkan perbaikan hanya pada sekumpulan kecil tugas penalaran yang relevan dengan mindware tersebut. Tidak akan ada langkah-langkah penalaran yang cukup untuk menguji hipotesis ini secara menyeluruh di sini, tetapi mungkin para pendukung TFD berharap agar perubahan domain-umum bertanggung jawab atas perbaikan apa pun sehingga matematika dapat dikatakan memiliki pengaruh yang lebih berguna dan tersebar luas pada keterampilan bernalar.

#### 4.2 RINGKASAN

Ada dua pertanyaan penelitian yang dibahas dalam bab ini: (a) apakah mempelajari matematika pada tingkat lanjutan terkait dengan peningkatan keterampilan penalaran? dan (b) jika ada peningkatan seperti itu, bagaimana mekanisme di baliknya? Pertanyaan-pertanyaan ini ditangani dengan studi longitudinal, di mana mahasiswa matematika tingkat AS dibandingkan dengan mahasiswa bahasa tingkat AS untuk pengembangan dalam penalaran logis dan keterampilan bernalar kritis. Para peserta diuji dua kali, pada awal dan akhir tahun studi AS mereka.

Tugas Inferensi Kondisional digunakan sebagai ukuran keterampilan penalaran logis dan tugas Silogisme Bias Keyakinan digunakan sebagai ukuran salah satu aspek bernalar kritis. Ukuran kecerdasan (RAPM) dan disposisi bernalar (CRT dan NFC) dimasukkan untuk menunjukkan apakah faktor umum domain pada tingkat kognisi algoritmik atau reflektif (Stanovich, 2009a) adalah mekanisme untuk pengembangan apa pun yang ditemukan. Langkah-langkah ini juga memungkinkan perbedaan yang sudah ada sebelumnya antara kelompok, yang mungkin ada karena desain eksperimen semu (lihat Bab 2), untuk dikontrol secara statistik. Akhirnya, tes matematika digunakan sebagai pemeriksaan manipulasi (untuk memastikan bahwa mahasiswa matematika memang belajar matematika lebih banyak daripada mahasiswa bahasa).

Sebelum menjelaskan studi longitudinal secara lebih rinci, tiga studi percontohan disajikan. Yang pertama menguji apakah tugas Belief Bias Silogisme dapat dibagi dua untuk menghemat waktu pengujian. Yang kedua menyelidiki apakah CRT dapat dimasukkan dalam masalah kata matematika, demi membuat sifat 'trik' dari pertanyaan-pertanyaan tersebut kurang mudah diingat tanpa mengubah cara peserta menanggapi. Studi percontohan akhir menilai durasi dan kesulitan langkah-langkah yang dipilih untuk memastikan kesesuaian untuk penelitian.

### 4.3 STUDI PILOT

#### Pilot 1 : Pebagian Tugas Silogisme

Dalam tugas *Silogisme Bias Keyakinan*, 24 silogisme disajikan dalam format yang dikontekstualisasikan. Dalam delapan silogisme, kepercayaan dan validitas sebelumnya sesuai (empat dipercaya-valid, empat sulit dipercaya-tidak valid), di delapan lainnya bertentangan (empat dipercaya-tidak valid, empat sulit dipercaya), dan di delapan terakhir konteksnya adalah netral (empat netral-valid, empat netral-tidak valid). Ini menciptakan empat masalah untuk masing-masing dari enam kombinasi logika-kepercayaan (lihat Gambar 4.1 untuk contoh masing-masing jenis). Dua dari masing-masing empat ini membuat pernyataan positif (P, Q), dan dua dari empat membuat pernyataan negatif (bukan-P, bukan-Q). Oleh karena itu, ada dua belas kombinasi kepercayaan, validitas, dan valensi. Dua soal untuk masing-masing kombinasi ini membuat total 24 item. Karena ada dua masalah dari setiap bentuk, tes dapat dibagi dua dan masih mencakup semua kombinasi untuk memberikan ukuran penuh bias keyakinan.

Tes dapat dibagi dengan cara ini sehingga setengah dari masalah dapat diberikan kepada peserta pada Waktu 1 dan setengah lainnya pada Waktu 2 (dengan urutan yang diseimbangkan oleh sekolah), untuk mengurangi efek pengujian ulang dan waktu pengujian. Metode ini berarti bahwa setiap perbedaan antara dua bagian dalam hal kepercayaan dapat menyebabkan keuntungan atau kerugian yang menyesatkan dalam bias kepercayaan antara dua titik waktu. Oleh karena itu, studi percontohan diperlukan untuk menentukan apakah ada perbedaan seperti itu. Dalam studi percontohan kesimpulan dari setiap masalah, di mana konflik kepercayaan/validitas berada, dinilai oleh peserta dalam hal seberapa dipercaya mereka. Soal dengan format yang sama dari setiap babak tes (misalnya soal yang dapat dipercaya dari Babak 1 dan Babak 2) kemudian dibandingkan untuk peringkat keterpercayaan.

#### 4.3.1 Metode

Peserta Lima puluh delapan peserta (38 laki-laki, usia 19-23, M=20.12) direkrut melalui email melalui tutor modul matematika dan mengambil bagian tanpa bayaran selama studi online yang lebih besar. Semua peserta adalah mahasiswa sarjana matematika dan teknik di Stekom University.

Prosedur Peserta mengambil bagian selama studi online yang tidak terkait tentang pilihan studi program gelar mereka (lihat Inglis, Palipana, Trenholm & Ward, 2011). Setelah menyelesaikan semua bagian yang relevan dengan pilihan studi mereka, mereka diminta untuk menyelesaikan satu bagian tentang keterpercayaan kalimat. Instruksinya berbunyi: "Di bawah ini adalah daftar kalimat. Beberapa kalimat akan benar-benar dapat dipercaya, beberapa akan benar-benar tidak dapat dipercaya, beberapa akan berada di tengah, dan beberapa tidak berarti. Tugas Anda adalah memutuskan yang mana".

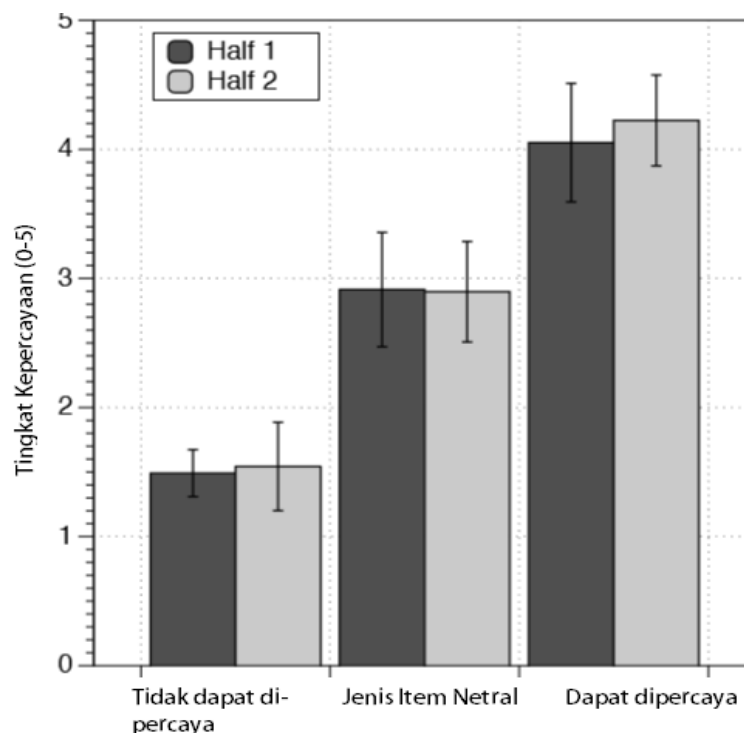
Di bawah instruksi, pada halaman yang sama, 24 kesimpulan dari tugas Silogisme Bias Keyakinan (lihat Lampiran C) disajikan dalam urutan yang ditetapkan, bergantian antara kesimpulan Setengah 1 dan kesimpulan Setengah 2. Di sebelah setiap kalimat terdapat skala 5 poin dengan opsi 'Sangat tidak dapat dipercaya', 'Cukup tidak dapat dipercaya', 'Tidak dapat dipercaya atau tidak dapat dipercaya', 'Cukup dapat dipercaya', dan 'Sangat dapat dipercaya'. Peserta menilai setiap pernyataan pada skala sebelum mengirimkan jawaban mereka.

#### 4.3.2 Hasil

24 kesimpulan silogisme tersebut dibagi menjadi enam kategori untuk analisis: Setengah 1 tidak dapat dipercaya (H1U), setengah 1 netral (H1N), setengah 1 dapat dipercaya

(H1B), setengah 2 tidak dapat dipercaya (H2U), setengah 2 netral (H2N), dan setengah 2 dipercaya (H2B). Tujuan dari analisis ini adalah untuk menguji perbedaan antara bagian tes pada peringkat kepercayaan setiap jenis item, yaitu, apakah kesimpulan H1U dinilai berbeda dengan kesimpulan H2U?

Respon rata-rata peserta ditunjukkan pada Gambar 4.2. Sebuah 2 (uji setengah: 1 atau 2)  $\times$  3 (kepercayaan dimaksudkan: dipercaya, netral, dipercaya) tindakan berulang analisis varians (ANOVA) dilakukan dengan peringkat kepercayaan sebagai variabel dependen. Ada efek utama yang signifikan dari kepercayaan yang dimaksudkan pada peringkat kepercayaan,  $F(2.114) = 323.2, p < .001, \eta^2 = .85$ . Tes post hoc menunjukkan bahwa item yang dapat dipercaya dinilai secara signifikan lebih dapat dipercaya ( $M = 4.14, SD = 0.50$ ) daripada item yang netral ( $M = 2.90, SD = 0.57$ ), yang pada gilirannya dinilai secara signifikan lebih dapat dipercaya daripada item yang tidak dapat dipercaya ( $M = 1.50, SD = 0.51$ ). Tidak ada efek utama yang signifikan dari uji-setengah, yang berarti bahwa tidak ada bukti dari dua bagian dari tes yang berbeda dalam kepercayaan. Yang penting, juga tidak ada interaksi yang signifikan antara setengah dan kepercayaan ( $p = .20, \eta^2 = .03$ ). H1U ( $M = 1.49, SD = 0.57$ ), H1B ( $M = 4.05, SD = 0.69$ ) dan H1N ( $M = 2.91, SD = 0.59$ ), kesimpulan dinilai mirip dengan H2U ( $M = 1.54, SD = 0.58$ ), H2B ( $M = 4.22, SD = 0.55$ ) dan H2N ( $M = 2.90, SD = 0.58$ ) kesimpulan masing-masing, sehingga tidak ada bukti bahwa kesimpulan di setiap setengah tes yang memiliki status kepercayaan yang sama berbeda dalam penilaian tingkat kepercayaan (Gambar 4.2).



**Gambar 4.2:** Peringkat kepercayaan untuk setiap jenis masalah dengan setengah tes (bilah kesalahan mewakili  $\pm 1$  SE dari rata-rata).

#### 4.3.3 Diskusi dan Implikasi

Ada tiga temuan dari penelitian ini: 1) masing-masing setengah dari tes dapat diasumsikan sama-sama dipercaya (setidaknya, tidak ada bukti perbedaan), 2) masalah kepercayaan yang sama dimaksudkan di setiap setengah dapat diasumsikan sama-sama dapat

dipercaya (sekali lagi, sejauh tidak ada bukti perbedaan), dan 3) kepercayaan yang dimaksudkan dari masalah atas tes secara keseluruhan sesuai dengan persepsi kepercayaan peserta. Ketiga hasil ini positif untuk penggunaan bagian yang berbeda dari tes pada titik waktu yang berbeda. Hasilnya menunjukkan bahwa seorang peserta akan menunjukkan tingkat bias keyakinan yang konsisten di setiap setengah tes, dan setiap perbedaan dari waktu ke waktu yang ditemukan dalam studi longitudinal tidak akan disebabkan oleh perbedaan dalam item yang digunakan tetapi karena perbedaan nyata pada peserta. 'kerentanan terhadap bias keyakinan.

### **Pilot 2 : Menyamarkan Tes Refleksi Kognitif**

Sejauh pengetahuan saya, CRT belum pernah digunakan dalam studi longitudinal sebelumnya. Perhatian khusus adalah sifat 'trik' dari pertanyaan mungkin membuat mereka lebih mudah diingat daripada ukuran lain di post-test dan bahwa ini akan mempengaruhi cara peserta (setidaknya mereka yang menghambat respons intuitif mereka dan dengan demikian menyadari 'trik') menanggapi mereka.

Upaya untuk mengatasi masalah ini adalah dengan mencampur pertanyaan CRT dengan masalah kata matematika non-trik, sehingga triknya mungkin kurang menonjol pada titik pengujian berulang. Namun, ada kemungkinan bahwa ini dapat mengubah sifat tes dalam beberapa cara, mungkin berarti peserta tidak lagi memproses pertanyaan CRT dengan cara yang seharusnya mereka lakukan. Tujuan dari studi percontohan yang dilaporkan di sini adalah untuk menilai apakah ini masalahnya. Peserta melihat tiga pertanyaan CRT diikuti oleh tiga masalah non-trik atau mereka melihat enam item dalam urutan acak. Tiga pertanyaan non-trik, yang ditunjukkan pada Gambar 5.3, diambil dari subtes Masalah Terapan Woodcock-Johnson III dan dipilih karena sederhana secara matematis dan memiliki panjang yang mirip dengan pertanyaan CRT. Jika pencampuran pertanyaan tidak mengubah sifat tes maka skor tidak boleh berbeda secara signifikan antar kelompok.

#### **4.3.4 Metode**

Peserta Peserta direkrut, tanpa pembayaran, melalui situs web yang mengiklankan studi penelitian berbasis internet di mana mereka melihat deskripsi singkat dan dapat membuka halaman web studi. Lima puluh empat peserta menyelesaikan semua enam pertanyaan dan dimasukkan dalam analisis. Para peserta berusia 18-59 ( $M=29,60$ ,  $SD=10,40$ ), dan 23 adalah laki-laki dan 31 perempuan. Dua puluh tujuh secara acak dialokasikan untuk kondisi campuran dan 27 untuk kondisi non-campuran ketika mereka membuka halaman web.

1. Jika seorang gadis menabung 1 setiap minggu selama 1 tahun, berapa banyak uang yang dia miliki pada akhir tahun itu?
  2. Jika seekor anjing dapat berlari sejauh dua seperempat mil dalam satu jam, berapa lama waktu yang dibutuhkan anjing tersebut untuk berlari sejauh empat setengah mil dengan kecepatan yang sama?
  3. Jarak tempuh bervariasi dari mobil ke mobil. Mobil Judy mendapat 22 mil untuk satu galon bensin, dan Bob mendapat 35 mil untuk satu galon gas. Berapa mil yang bisa Judy kendarai dengan enam galon bensin?
- Jawaban: Q1 = 52, Q2 = 2 jam, Q3 = 132 mil

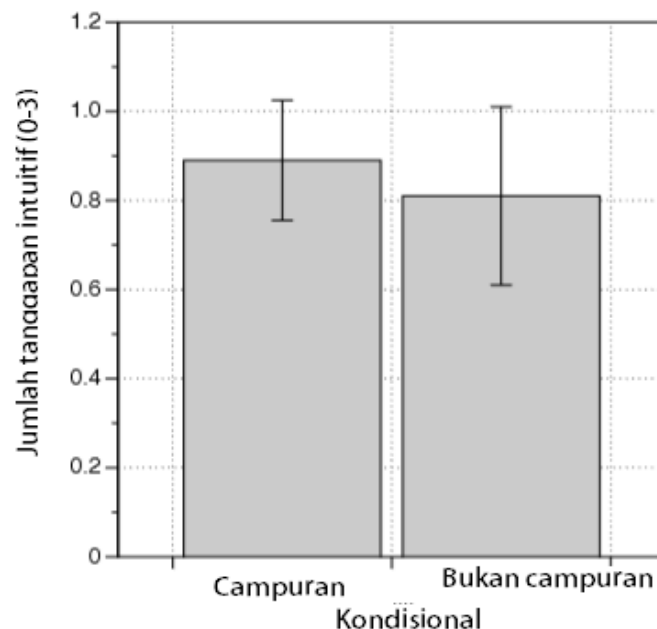
**Gambar 4.3:** Tiga item dari subtes Masalah Terapan Woodcock Johnson III yang digunakan dalam Studi Percontohan 2.

Prosedur Peserta pertama kali melihat halaman yang menyediakan informasi tentang penelitian. Mereka diberitahu bahwa jika mereka mengambil bagian mereka akan diminta untuk menjawab enam soal aritmatika kata yang akan memakan waktu tidak lebih dari 10 menit. Mereka juga diberitahu bahwa data mereka akan dijaga kerahasiaannya dan hanya digunakan untuk tujuan penelitian. Mereka diminta untuk memilih apakah mereka ingin berpartisipasi secara serius atau hanya menelusuri halaman sebelum melanjutkan studi, dan hanya mereka yang ingin berpartisipasi secara serius yang dimasukkan dalam analisis.

Selanjutnya mereka diminta untuk melaporkan jenis kelamin, usia, jurusan (jika ada), dan apakah bahasa ibu mereka adalah bahasa atau non-Indonesia. Keenam pertanyaan masing-masing disajikan pada halaman terpisah dan mengharuskan peserta untuk mengetik jawaban mereka ke dalam kotak jawaban kosong. Dalam kondisi campuran, enam pertanyaan disajikan secara acak. Dalam kondisi non-campuran, tiga pertanyaan CRT disajikan pertama dalam urutan acak, diikuti oleh tiga pertanyaan nontrik dalam urutan acak. Akhirnya, peserta mengucapkan terima kasih atas partisipasi mereka dan diberikan alamat email saya jika mereka ingin meminta informasi lebih lanjut atau untuk menarik data mereka (tidak ada).

#### 4.3.5 Hasil

Dua tes Mann-Whitney U digunakan untuk membandingkan kinerja di dua kondisi: satu menganalisis jumlah tanggapan yang benar untuk tiga pertanyaan CRT, dan satu menganalisis jumlah tanggapan intuitif (mungkin, tetapi jarang, untuk memberikan jawaban non-intuitif tetapi tanggapan yang salah yang berarti bahwa ukuran-ukuran ini bukan kebalikan yang tepat satu sama lain).



**Gambar 4.4:** Jumlah rata-rata respons intuitif dalam kondisi campuran dan non-campuran (bilah kesalahan mewakili  $\pm 1$  SE dari rata-rata).

Jumlah jawaban yang benar tidak berbeda nyata pada kelompok campuran ( $M=2.00$ ,  $SD=0.73$ ) dan non-campuran ( $M=1.93$ ,  $SD=1.07$ ),  $U(54)=358.0$ ,  $z = .12$ ,  $p = 0,904$ ,  $r = 0,02$ . Demikian pula, jumlah tanggapan intuitif tidak berbeda secara signifikan pada kelompok campuran ( $M=0.89$ ,  $SD=0.70$ ) dan non-campuran ( $M=0.81$ ,  $SD=1.04$ ),  $U(54)=417.00$ ,  $z = .98$ ,  $p=.325$ ,  $r = 0.13$ , lihat Gambar 5.4.



#### 4.3.6 *Diskusi dan Implikasi*

Tujuan dari studi percontohan ini adalah untuk menilai apakah pencampuran tiga pertanyaan CRT dengan tiga pertanyaan non-trik akan mempengaruhi cara peserta menanggapi pertanyaan CRT. Tidak ada bukti bahwa ini masalahnya. Baik jumlah respons yang benar maupun jumlah respons intuitif yang diberikan pada pertanyaan CRT tidak berbeda secara signifikan ketika digabungkan dengan pertanyaan non-trik dibandingkan saat pertama kali dilihat.

Satu hal yang disarankan oleh hasil ini adalah level reflektif menentukan berdasarkan item demi item apakah level algoritmik harus digunakan. Ini bukanlah kasus bahwa menjawab item tanpa jawaban intuitif yang dengan demikian memerlukan pemrosesan tingkat algoritmik (salah satu pertanyaan Woodcock-Johnson) menetapkan tingkat reflektif ke pola pikir 'gunakan pemrosesan algoritmik'. Jika ini masalahnya, kehadiran item yang memerlukan pemrosesan algoritmik berarti bahwa jawaban intuitif terhambat untuk pertanyaan CRT berikutnya, yang tidak terjadi di sini. Ini sesuai dengan temuan yang sudah mapan bahwa peserta tidak cenderung menjawab semua pertanyaan CRT dengan benar atau salah, melainkan mereka mungkin mendapatkan satu atau dua yang benar dan memberikan tanggapan intuitif kepada yang lainnya (Frederick, 2005).

Kembali ke tujuan studi percontohan ini, hasil menunjukkan bahwa tiga pertanyaan CRT dan tiga pertanyaan Woodcock-Johnson dapat dicampur secara acak dalam studi longitudinal tanpa mempengaruhi cara peserta menanggapi pertanyaan CRT, yang merupakan ukuran sebenarnya (tanggapan atas pertanyaan Woodcock-Johnson tidak akan dianalisis). Pertanyaan Woodcock-Johnson tampaknya cukup sederhana sehingga tidak mempengaruhi tingkat reflektif secara umum. Diharapkan inklusi mereka akan mengurangi efek pengujian ulang pada kinerja CRT.

#### **Pilot 3 : Durasi dan Pengukuran Kesulitan**

Ada beberapa ukuran yang dipilih untuk studi longitudinal yang dilaporkan di bawah ini yang waktu penyelesaian rata-ratanya tidak diketahui. Untuk merencanakan slot waktu yang sesuai dengan sekolah, studi percontohan dilakukan untuk menentukan total durasi sesi yang diperlukan. Lima mahasiswa matematika sarjana direkrut untuk menyelesaikan seluruh rangkaian tugas (demografi, subset RAPM, Tugas Inferensi Kondisional, tugas Silogisme Bias Keyakinan, CRT, skala Kebutuhan Kognisi, dan tugas matematika). Tujuannya adalah untuk mencatat durasi setiap pengukuran individu dan durasi keseluruhan tes untuk setiap peserta, sehingga durasi rata-rata dapat diturunkan. Tujuan selanjutnya adalah untuk menilai apakah langkah-langkah yang digunakan memiliki kesulitan yang sesuai dan apakah instruksi untuk setiap ukuran jelas bagi peserta yang tidak terbiasa dengan tugas-tugas tersebut.

#### 4.3.7 *Metode*

Peserta Para peserta adalah lima mahasiswa sarjana matematika (satu laki-laki, empat perempuan), berusia 19 sampai 51 ( $L=25.80$ ,  $SD=14.10$ ) yang mengambil bagian dengan imbalan masing-masing Rp. 150.000. Rekrutmen dilakukan melalui iklan email kepada mahasiswa pada modul persamaan diferensial untuk mahasiswa tahun pertama dan kedua. Sampel ini diasumsikan memiliki kemampuan umum yang lebih tinggi daripada mahasiswa tingkat AS dari studi longitudinal, meskipun tidak terlalu banyak. Agaknya cenderung mahasiswa tingkat AS yang paling mampu yang melanjutkan ke studi tingkat sarjana, sehingga sampel sarjana di sini mungkin mirip dengan mahasiswa yang lebih mampu dalam sampel tingkat AS. Implikasinya adalah skor yang ditemukan di sini mungkin sedikit lebih tinggi

daripada di sampel tingkat AS, sehingga setiap efek rantai yang ditemukan harus menjadi perhatian khusus.

Prosedur. Peserta diberitahu bahwa mereka akan diberikan buku tes untuk diselesaikan. Mereka diberitahu bahwa tujuan dari penelitian ini adalah untuk menentukan lamanya tes, yang akan digunakan dalam penelitian skala besar di masa depan. Mereka kemudian diminta untuk menandatangani formulir persetujuan sebelum mengambil bagian. Dua peserta mengambil bagian secara bersamaan tetapi bekerja sendiri, dan tiga mengambil bagian secara individu. Semua pengujian berlangsung di ruang seminar yang tenang. Peserta diminta untuk mengerjakan buklet dengan kecepatan mereka sendiri, memberi tahu eksperimen ketika mereka mencapai akhir bagian dan memulai bagian berikutnya. Bagian disajikan dalam urutan yang ditetapkan untuk semua peserta. Setelah mereka menyelesaikan buklet, peserta ditanya apakah ada bagian dari tes yang tidak jelas, terlalu mudah atau terlalu sulit, dan apakah mereka memiliki komentar lain. Mereka kemudian berterima kasih, dibayar, dan diberhentikan.

#### 4.3.8 Hasil

Bagian pertama dari hasil akan membahas panjang tindakan, bagian kedua akan membahas kisaran skor yang diperoleh, dan bagian terakhir akan membahas komentar peserta tentang kejelasan tugas.

**Durasi.** Tabel 4.1 menunjukkan statistik deskriptif untuk lamanya waktu yang dibutuhkan untuk setiap pengukuran. Durasi total rata-rata adalah 45,40 menit dengan standar deviasi 11,63 menit. Namun, total data durasi miring positif (2.13) dengan empat titik data pada rentang 39-43 menit dan satu titik data 66 menit. Oleh karena itu, empat dari lima peserta menyelesaikan tes lebih cepat dari waktu rata-rata 45,4 menit. Bagian RAPM dari tes memiliki batas waktu 15 menit. Namun, dapat dilihat dari tabel bahwa beberapa peserta menyelesaikan lebih cepat dari ini.

**Skor.** Nilai rata-rata diperiksa untuk menunjukkan apakah ada tugas yang terkena efek rantai atau langit-langit. Tabel 4.2 menunjukkan statistik deskriptif untuk skor yang diperoleh pada setiap ukuran serta skor minimum dan maksimum teoritis. Skor CRT tidak diperiksa karena peserta baru-baru ini dihadapkan pada tugas sebagai bagian dari studi lain. Seperti yang ditunjukkan tabel, tidak ada skor tugas rata-rata yang berada pada, atau mendekati, skor minimum atau maksimum teoretis, dengan kemungkinan pengecualian dari tugas Silogisme Bias Keyakinan yang mendekati maksimum teoretis.

Komentar tambahan. Semua lima peserta berkomentar bahwa Tugas Inferensi Kondisional tidak sepenuhnya jelas pada pembacaan pertama dari instruksi, meskipun mereka menemukan bahwa itu menjadi jelas ketika mereka mulai menyelesaikannya (dibahas di bawah).

**Tabel 4.1:** Informasi durasi untuk setiap ukuran yang digunakan dalam buku tes (satuan dalam menit).

<i>Pengukuran</i>	<i>Mean</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Demografis	2.2	1.64	1	5
Matrik Raven	14.0	1.73	11	15
Interface Kondisional	10.6	3.13	8	16
Silogisme Bias Kepercayaan	3.8	1.3	3	6
Tes refleksi Kognitif	1.2	0.45	1	2

Butuh Cognisi	3.2	1.79	2	6
Matematika	10.4	3.64	7	16
Total	45.4	11.63	39	66

#### 4.3.9 Diskusi dan implikasi

Hasil uji coba menunjukkan bahwa (a) durasi rata-rata dari keseluruhan tes adalah 45 menit, dengan mayoritas peserta menyelesaikan dalam waktu yang lebih singkat, (b) langkah-langkah di atas pada beberapa individu tetapi tidak rata-rata, dan (c) ada satu masalah yang dicatat oleh peserta; instruksi untuk Tugas Inferensi Kondisional.

Lamanya waktu yang ditempuh berarti seluruh tes dapat diselesaikan dalam satu pelajaran sekolah, yang biasanya berdurasi 50-60 menit. Membutuhkan hanya satu sesi akan mengurangi permintaan waktu guru dan peserta. Di sisi lain, terbukti dari satu peserta yang mengambil 66 menit bahwa beberapa individu mungkin membutuhkan waktu lebih lama dari satu pelajaran, dalam hal ini mereka harus meninggalkan tes tidak lengkap, tinggal setelah pelajaran selesai, atau lanjutkan di lain waktu. Ini adalah sesuatu yang dapat diputuskan dengan masukan dari guru yang bersangkutan.

**Tabel 4.2:** Statistik deskriptif untuk setiap ukuran yang digunakan dalam buku tes (standar deviasi dalam tanda kurung).

<i>Pengukuran</i>	<i>Mean (SD)</i>	<i>Kemungkinan Min</i>	<i>Observasi Min</i>	<i>Kemungkinan Max</i>	<i>Observasi Max</i>
	11.40				
Metrik Raven	(3.78)	0	8	18	16
Kondisional/Kondisional	22.40				
	(7.23)	0	15	32	32
Silogisme	10.60				
	(1.67)	0	8	12	12
NFC	5.47				
	(1.09)	0	4.56	9	7.33
Matematika	10.20				
	(2.49)	0	7	15	12

Dengan mengacu pada bagian kedua dari hasil, skor yang diperoleh hanya bermasalah untuk tugas Silogisme Bias Keyakinan, di mana ada sedikit efek plafon. Perlu ditekankan kembali bahwa peserta dalam percontohan ini secara pendidikan lebih maju daripada peserta studi utama yang dilaporkan di bawah ini. Studi utama akan menggunakan peserta dari awal hingga akhir tahun studi level AS mereka, sedangkan peserta di sini adalah pada akhir tahun pertama atau kedua gelar sarjana matematika, sehingga dapat diharapkan prestasi mereka lebih tinggi. tes dari mayoritas peserta studi utama akan. Oleh karena itu, semua langkah yang diujicobakan di sini diharapkan memberikan variasi yang cukup untuk mendeteksi peningkatan selama level AS.

Tugas Inferensi Kondisional yang dicatat oleh semua peserta mengatakan bahwa instruksi tidak sepenuhnya jelas. Namun, tidak ada peserta yang dapat menyarankan bagaimana instruksi dapat diklarifikasi bahkan setelah mereka menyelesaikan tes dan melaporkan bahwa mereka memahaminya. Ini mungkin mencerminkan sifat tugas yang sulit dihindari, karena ini bukan sesuatu yang biasanya ditemui dalam kehidupan sehari-hari.

Instruksi yang digunakan diadaptasi dari Evans, Clibbens & Rood (1995), yang tidak melaporkan masalah serupa dalam penggunaan ukuran skala besar mereka.

Demi konsistensi dengan penelitian yang dipublikasikan, instruksi akan dibuat identik dengan yang digunakan oleh Evans, Clibbens & Rood untuk studi utama, dan diharapkan bahkan jika para peserta menemukan instruksi yang rumit dalam isolasi, tugas akan menjadi jelas. begitu mereka mulai. Eksperimen akan selalu hadir ketika peserta menyelesaikan tugas-tugas ini, sehingga akan ada kesempatan untuk meminta klarifikasi jika diperlukan. Singkatnya, studi percontohan yang dijelaskan di sini tidak menimbulkan masalah yang memerlukan tindakan yang dipilih untuk diubah atau diganti.

#### **4.4 RINGKASAN STUDI PILOT**

Tiga studi percontohan telah disajikan yang menilai berbagai aspek dari langkah-langkah yang dipilih untuk studi utama, dan masing-masing telah memberikan hasil yang positif. Selanjutnya, studi utama itu sendiri disajikan.

#### **4.5 STUDI UTAMA**

Untuk rekap, ada dua pertanyaan penelitian untuk studi utama:

1. Apakah belajar matematika di tingkat AS ada pengaruhnya dengan peningkatan keterampilan penalaran?
2. Jika ada peningkatan seperti itu, bagaimana mekanisme di baliknya?

#### **4.6 METODE**

##### **4.6.1 Desain**

Penelitian ini mengikuti desain kuasi-eksperimental longitudinal. Peserta direkrut setelah mereka memilih mata pelajaran tingkat AS mereka dan diuji pada awal (selama semester pertama dan sedekat mungkin dengan awal semester) dan akhir (setelah pengajaran selesai) tahun studi AS mereka. Mereka menyelesaikan serangkaian tugas yang sama pada kedua titik waktu (dengan pengecualian tugas Silogisme Bias Keyakinan seperti yang dijelaskan dalam Percontohan 1).

##### **4.6.2 Peserta**

Seratus dua puluh empat peserta direkrut dari lima Universitas di Leicestershire, Hampshire dan Derbyshire. Tujuh puluh tujuh mempelajari matematika tingkat AS atau matematika tingkat lanjut di antara mata kuliah lain dan empat puluh tujuh mempelajari bahasa AS dan bukan matematika. Mahasiswa bahasa berperan sebagai kelompok pembanding. Peserta dan orang tua/wali memberikan persetujuan tertulis.

Dari sampel asli, 44 mahasiswa matematika dan 38 mahasiswa bahasa mengambil bagian pada kedua titik waktu tersebut. Tidak ada perbedaan pada salah satu ukuran waktu 1 antara mereka yang kembali dan mereka yang tidak ( $p > .15$ ). Kelompok matematika terdiri dari 21 perempuan dan 23 laki-laki dan kelompok bahasa terdiri dari 23 perempuan dan 15 laki-laki. Delapan puluh peserta melaporkan bahasa pertama mereka sebagai bahasa, satu melaporkan bahasa Gujarati dan Perancis, dan satu tidak melaporkan bahasa pertama mereka. Empat peserta dalam kelompok matematika melaporkan telah didiagnosis dengan disleksia dan dua di kelompok bahasa melaporkan bahwa mereka diduga menderita disleksia. Tidak ada yang melaporkan mengalami diskalkulia atau disabilitas lain yang relevan.

Peserta juga melaporkan nilai GCSE (Sertifikat Umum Pendidikan Menengah, ujian akhir wajib belajar) mereka. Dari sini, skor pencapaian sebelumnya dihitung untuk setiap peserta untuk digunakan sebagai kovariat dalam analisis yang dilaporkan di bawah ini. Ini adalah jumlah nilai yang dicapai, dengan nilai A\* diberi nilai 8, nilai A diberi nilai 7, nilai B diberi nilai 6, dan seterusnya. Skor dijumlahkan daripada dirata-rata untuk memperhitungkan variasi jumlah GCSE yang diambil – mahasiswa yang mencapai 10 GCSE di kelas A dapat dikatakan memiliki tingkat pencapaian yang lebih tinggi daripada mahasiswa yang mencapai 9 GCSE di kelas A, misalnya, dari mereka yang belajar matematika tingkat AS, 25 telah mencapai A\* di matematika GCSE, 15 mencapai A, 3 mencapai B dan 1 tidak melaporkan nilai mereka. Dari mereka yang belajar bahasa, ada 5 A\*s dalam matematika GCSE, 6 As, 14 Bs, 11 Cs, 1 D dan 1 E.

#### 4.6.3 *Silabus Matematika*

Ada tiga versi berbeda dari kuliah matematika tingkat AS yang tersedia untuk mahasiswa, semuanya memiliki konten yang serupa. **Silabus** berisi bagian tentang aljabar, geometri koordinat, kalkulus pengantar, trigonometri, probabilitas, pemodelan matematika, kinematika, hukum gerak Newton, dan gaya (misalnya, Asesmen dan Aliansi Kualifikasi, 2011), di antara topik lainnya. Kuliah ini dianggap oleh beberapa orang sebagai persiapan yang cukup mendasar dan tidak memadai untuk studi matematika tingkat universitas (misalnya Lawson, 1997, 2003), dan yang paling penting, mahasiswa tidak diajarkan matematika berbasis bukti, atau definisi kondisional. Hal ini dikonfirmasi dengan analisis setiap ujian matematika tingkat AS antara 2009 dan 2011. Dari 929 pertanyaan yang ditetapkan, hanya satu yang berisi kalimat 'jika kemudian' yang eksplisit, dan tidak ada penyebutan istilah 'modus ponens', 'modus tollens' atau 'kondisional'.

#### 4.7 HASIL PENGUKURAN

Inferensi Kondisional. Peserta menyelesaikan standar 32 item Tugas Inferensi Kondisional (Evans et al., 1995), yang terdiri dari delapan item masing-masing dari empat jenis inferensi: modus ponens (MP), penolakan anteseden (DA), penegasan konsekuensi (AC) dan Modus Tollens (MT). Setengah dari item menggunakan negasi eksplisit (misalnya "bukan 5") dan setengahnya menggunakan negasi implisit (misalnya "bukan 5" direpresentasikan sebagai, misalnya, 6). Jenis inferensi yang digunakan diringkas dalam Tabel 5.3 dan ukuran lengkapnya disajikan dalam Lampiran B.

Konten leksikal dihasilkan secara acak dan item disajikan dalam urutan acak untuk setiap peserta. Peserta memutuskan apakah setiap item itu valid (yaitu kesimpulan harus diikuti, dengan asumsi bahwa premis itu benar) atau tidak valid. Enam langkah diambil:

1. Indeks kondisional material (MCI, jumlah jawaban dari 32 yang konsisten dengan interpretasi materi), yang dihitung sebagai: jumlah kesimpulan MP yang didukung + (8–jumlah kesimpulan DA yang didukung) + (8–jumlah kesimpulan AC yang didukung) + jumlah inferensi MT yang didukung.
2. Indeks kondisional yang rusak (DCI, jumlah jawaban dari 32 yang konsisten dengan interpretasi yang salah), yang dihitung sebagai: jumlah inferensi MP yang didukung + (8–jumlah inferensi DA yang didukung) + (8–jumlah inferensi AC didukung) + (8–jumlah kesimpulan MT yang didukung).
3. Indeks bikondisional (BCI, jumlah jawaban dari 32 yang konsisten dengan interpretasi bikondisional), yang dihitung sebagai: jumlah inferensi MP yang didukung + jumlah

inferensi DA yang didukung + jumlah inferensi AC yang didukung + jumlah inferensi MT yang didukung.

4. Indeks kondisional konjungtif (CCI, jumlah jawaban dari 32 yang konsisten dengan interpretasi konjungtif), yang dihitung sebagai: jumlah kesimpulan MP yang didukung + (8–jumlah kesimpulan DA yang didukung) + jumlah kesimpulan AC yang didukung + (8– jumlah inferensi MT yang didukung).
5. Indeks kesimpulan negatif (NCI), yang dihitung sebagai jumlah kesimpulan yang didukung pada argumen dengan kesimpulan negatif dikurangi jumlah kesimpulan yang didukung pada argumen dengan kesimpulan afirmatif.
6. Indeks premis afirmatif (API), yang dihitung sebagai jumlah inferensi yang didukung pada argumen dengan premis afirmatif dikurangi jumlah inferensi yang didukung pada argumen dengan premis negatif.

Instruksi yang diberikan identik dengan yang digunakan oleh Evans et al. (1995). Contoh item ditunjukkan pada Gambar 5.5. Silogisme. Tugas Silogisme Bias Keyakinan (disajikan secara lengkap dalam Lampiran C) digunakan sebagai ukuran kemampuan untuk bernalar secara independen dari keyakinan sebelumnya (Evans et al., 1983; Markovits & Nantel, 1989; S'a et al., 1999) . Tugas terdiri dari 12 silogisme kontekstual, empat kongruen (dapat dipercaya-valid, luar biasa-tidak valid), empat tidak kongruen (dapat dipercaya-tidak valid, luar biasa-valid) dan empat netral (contoh item ditunjukkan pada Gambar 5.1). Peserta memutuskan apakah setiap silogisme secara logis valid atau tidak setelah diinstruksikan untuk mengabaikan keyakinan mereka sebelumnya. Dua ukuran diambil: skor total dari 12, yang menunjukkan kemampuan penalaran silogistik, dan Indeks Bias Keyakinan. Indeks Bias Keyakinan dihitung untuk setiap peserta dengan mengurangi jumlah item tidak kongruen yang dijawab dengan benar dari jumlah item kongruen yang dijawab dengan benar. Skor yang dihasilkan menunjukkan sejauh mana jawaban seseorang dipengaruhi oleh kepercayaan atau validitas. Indeks Bias Keyakinan dapat berkisar dari -4 hingga +4 dengan skor positif yang menunjukkan beberapa derajat bias keyakinan.

**Tabel 4.3:** Empat inferensi dan jenis pernyataan kondisional dengan dan tanpa premis negasi (Pr) dan kesimpulan (Con). Simbol harus dibaca 'tidak'.

Kondisional	MP		DA		AC		MT	
	Pr	Con	Pr	Con	Pr	Con	Pr	Con
jika p maka q	p	q	-p	-q	q	p	-q	-p
jika p maka -q	p	-q	-p	q	-q	p	q	-p
jika -p maka q	-p	q	p	-q	q	-p	-q	p
jika -p maka -q	-p	-q	p	q	-q	-p	q	p

Jika hurufnya H maka angkanya 5

Hurufnya H

Kesimpulan: Angkanya 5

Ya

Tidak

**Gambar 4.5:** Contoh item dari Tugas Inferensi Kondisional.

**Matriks Progresif Lanjutan Raven (RAPM).** Sebuah subset 18 item RAPM (lihat Lampiran A) dengan batas waktu 15 menit (S'a et al., 1999) digunakan sebagai ukuran kecerdasan umum (atau pemrosesan tingkat algoritmik, Stanovich, 2009a).

**Tes Refleksi Kognitif/Cognitive Reflective Test (CRT).** Seperti yang disarankan oleh Toplak et al. (2011) jumlah tanggapan intuitif yang diberikan pada CRT tiga item (Frederick, 2005, lihat Gambar 2.10) digunakan sebagai ukuran kinerja kecenderungan untuk menggunakan pemrosesan Tipe 2 (pada tingkat reflektif). Skor dibalik sehingga skor yang lebih tinggi mewakili kinerja yang lebih normatif, sejalan dengan ukuran lainnya. Pertanyaan-pertanyaan tersebut dicampur secara acak dengan tiga soal matematika sederhana dengan panjang yang sama dari subtes Masalah Terapan Woodcock-Johnson III seperti yang dijelaskan dalam bab 4.2.2.

**Kebutuhan akan Kognisi/Need for Cognitive(NFC).** Skala NFC (Cacioppo, Petty, Feinstein & Jarvis, 1996, lihat Lampiran D) dimasukkan sebagai ukuran laporan diri dari disposisi bernalar untuk melengkapi ukuran CRT berbasis kinerja. Namun, tidak ada perbedaan antara kelompok pada Waktu 1 ( $p = 0,616$ ) atau pada Waktu 2 ( $p = 0,374$ ) atau perubahan dari waktu ke waktu pada kedua kelompok (keduanya  $ps > 0,670$ ), meskipun ukuran berkorelasi secara signifikan pada Waktu 1 dengan skor RAPM  $r(122) = .19, p = .034$ , Skor silogisme,  $r(122) = .31, p = .001$ , Indeks Bias Keyakinan,  $r(122) = .30, p = .001$ , jawaban yang benar untuk CRT,  $r(122) = .19, p = .040$ , MCI,  $r(120) = .21, p = .022$ , dan DCI,  $r(122) = .21, p = 0,021$ . Karena kurangnya perbedaan antar kelompok atau perubahan dari waktu ke waktu, skor NFC tidak dibahas lebih lanjut.

**Pemeriksaan Manipulasi Matematika.** Sebuah tes matematika 15 item dimasukkan sebagai cek manipulasi. Hal ini untuk memastikan bahwa mahasiswa yang belajar matematika tingkat AS memang belajar matematika lebih banyak daripada mereka yang belajar bahasa tingkat AS dan bukan matematika. Dalam hal tidak ada peningkatan dalam keterampilan penalaran dalam kelompok matematika, akan berguna untuk mengesampingkan kemungkinan bahwa alasan mereka tidak meningkat adalah karena mereka tidak benar-benar meningkat dalam matematika.

Dua belas item diambil dari subtes Perhitungan Woodcock-Johnson III. Sembilan telah menunjukkan akurasi rata-rata kurang dari 55% dan berkorelasi dengan kinerja di seluruh tes di 0,86 dalam dataset sebelumnya dengan mahasiswa sarjana disiplin campuran (Inglis, Attridge, Batchelor & Gilmore, 2011) dan tiga sisanya diambil dari kisaran yang lebih rendah untuk mencegah efek lantai pada mahasiswa bahasa.

Tiga item terakhir adalah pertanyaan paling sulit pada tes diagnostik Universitas Stekom untuk sarjana matematika baru berdasarkan kinerja pada tahun 2008 dan 2009. Tes diagnostik dirancang untuk menguji kemampuan mahasiswa yang masuk dengan matematika A-level, dan tiga item dimasukkan untuk mencegah efek langit-langit pada mahasiswa matematika pada titik waktu kedua sambil memastikan bahwa konten tidak dikembangkan secara tidak tepat. Pertanyaan disajikan dalam urutan yang dimaksudkan untuk menjadi progresif. Tugas selengkapnya disajikan dalam Lampiran E.

#### 4.7.1 Prosedur

Peserta mengambil bagian dalam kelompok (4-34) selama hari sekolah di bawah kondisi gaya ujian. Semua tugas diberikan dalam satu buklet kertas. Tugas RAPM selalu diselesaikan terlebih dahulu agar batas waktu 15 menit dapat ditegakkan, dan urutan tugas-tugas berikutnya diseimbangkan antar-peserta mengikuti desain Latin Square. Peserta diinstruksikan untuk bekerja dengan kecepatan mereka sendiri sampai mereka menyelesaikan

semua tugas dan sesi berlangsung sekitar 45 menit. Di empat dari lima sekolah, peserta diikutsertakan dalam undian berhadiah untuk memenangkan Nintendo DS Lite atau pemutar DVD portabel pada Time 1. At Time 2, masing-masing peserta dibayar Rp. 48.000 untuk ambil bagian. Di sekolah kelima, guru lebih memilih mahasiswa untuk mengambil bagian tanpa insentif eksternal.

Hasilnya dilaporkan dalam tiga bagian:

- (i) analisis awal
- (ii) pengembangan skor *Inferensi Kondisional dan Silogisme*
- (iii) mekanisme pengembangan.

#### 4.7.2 Analisis awal

**Inklusi data.** Empat puluh empat mahasiswa matematika dan tiga puluh delapan mahasiswa bahasa mengambil bagian di kedua titik waktu dan dimasukkan dalam analisis. Mereka yang putus sekolah biasanya berganti jurusan; tidak ada perbedaan yang signifikan dalam skor Time 1 pada salah satu ukuran antara mereka yang mengambil bagian pada Time 2 dan mereka yang drop out ( $ps > .15$ ).

**Keandalan.** Keandalan Tugas Inferensi Kondisional dinilai dengan kumpulan data besar yang dikumpulkan dari beberapa penelitian (termasuk Inglis & Simpson, 2008, 2009a; Inglis, Attridge et al., 2011, dan studi yang dilaporkan dalam bab saat ini dan Bab 6 ini penelitian). Ini menghasilkan kumpulan data 656 peserta dari tiga universitas dan lima sekolah dan perguruan tinggi. Alpha Cronbach, 0,87, ditemukan cukup tinggi untuk ukuran yang dianggap andal secara internal.

**Keandalan tugas Silogisme Bias.** Keyakinan dinilai hanya dengan menggunakan data yang disajikan di sini. Alpha Cronbach ditemukan lebih rendah, pada 0,65, tetapi tidak terlalu rendah. Ini mungkin karena sifat dua jenis item yang sengaja dibuat berbeda – keyakinan/validitas yang konsisten dan sepertiga tugas yang tidak konsisten.

**Kovariat.** Kelompok matematika mendapat nilai yang lebih tinggi secara signifikan pada RAPM ( $M = 9.57, SD = 3.26$ ) dibandingkan kelompok bahasa pada Time 1 ( $M = 7.03, SD = 3.45$ ),  $t(80) = 3.43, p = .001, d = 0,76$ . Kelompok matematika juga mendapat skor yang lebih tinggi secara signifikan pada CRT (jumlah jawaban intuitif terbalik,  $M = 1,77, SD = 1,12$ ) dibandingkan kelompok bahasa pada Time 1 ( $M = 0,89, SD = 0,86$ ),  $U(82) = 466,00, z = 4,29, p < .001, r = .39$ . Skor pencapaian akademik sebelumnya berkisar antara 30 hingga 99 ( $M = 64.10, SD = 11,94$ ) dan sedikit lebih tinggi pada kelompok matematika ( $M = 66,26, SD = 9,75$ ) dibandingkan kelompok bahasa ( $M = 61,66, SD = 13,45$ ),  $t(79) = 1,75, p = 0,084, d = 0,39$ .

Skor MCI pada Time 1 secara signifikan berkorelasi dengan RAPM,  $r(79) = .41, p < .001$ , CRT,  $r(79) = .42, p < .001$ , dan skor pencapaian sebelumnya,  $r(78) = .30, p = .007$ . Akibatnya, RAPM, CRT dan skor pencapaian akademik sebelumnya digunakan sebagai kovariat dalam analisis selanjutnya dari skor Inferensi Kondisional.

Skor silogisme pada Time 1 juga berkorelasi signifikan dengan RAPM,  $r(82) = .43, p < .001$ , CRT,  $r(82) = .48, p < .001$ , dan skor pencapaian akademik sebelumnya,  $r(81) = .38, p < .001$ , mendukung penggunaan ketiga ukuran sebagai kovariat dalam analisis skor Silogisme.

Terakhir, skor Belief Bias Index (BBI) pada Time 1 juga berkorelasi dengan RAPM,  $r(82) = .28, p = .011$ , CRT,  $r(82) = .26, p = .019$ , dan sebelumnya skor pencapaian akademik,  $r(81) = .28, p = .010$ , sehingga kovariat juga digunakan dalam analisis BBI.

Meskipun kedua kelompok meningkatkan skor RAPM dan CRT mereka sedikit selama tahun ini, tidak ada efek interaksi Kelompok  $\times$  Waktu yang mendekati signifikan,  $ps > .20$ .

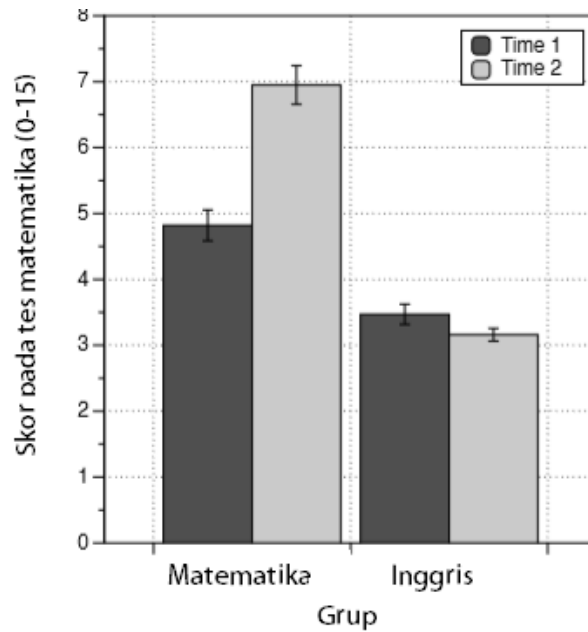


**Pemeriksaan Manipulasi.** Perubahan nilai tes matematika dianalisis dengan  $2 \times 2$  (Grup: matematika) Analisis Varians (ANOVA). Ada interaksi yang signifikan,  $F(1,80) = 52.91, p < .001, \eta^2 = .40$ , yang menunjukkan bahwa kelompok matematika meningkat ke tingkat yang lebih besar (Waktu 1  $M = 4.82, SD = 1.56$ , Waktu 2  $M = 6.95, SD = 1.94$ ) daripada kelompok bahasa (Time 1  $M = 3.47, SD = 0.95$ , Time 2  $M = 3.12, SD = 0.59$ , lihat Gambar 4.6). Peningkatan kelompok matematika dari waktu ke waktu dikonfirmasi oleh perbandingan yang direncanakan dari skor Waktu 1 dan 2,  $t(43) = 7.37, p < .001, d = 1.21$ . Ini menunjukkan bahwa sebagai kelompok mereka terlibat dan belajar dari tahun mereka belajar matematika dan manipulasi kuasi berhasil.

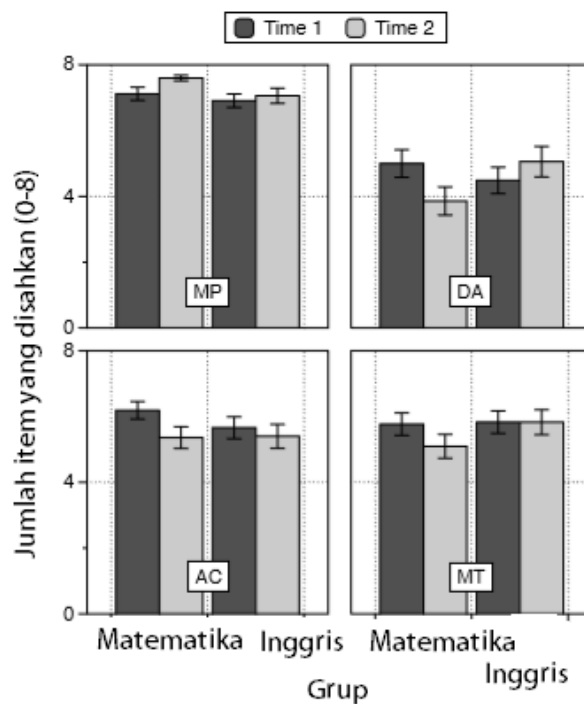
#### **4.7.3 Analisis Inferensi Kondisional.**

**Tingkat pengesahan.** Tingkat dukungan masing-masing kelompok pada Waktu 1 dianalisis dengan ANOVA  $2 \times 4$  dengan satu faktor dalam mata pelajaran: Jenis Inferensi (MP, DA, AC, MT), dan satu faktor antara mata pelajaran: Kelompok (matematika, Bahasa). Ada efek utama yang signifikan dari Jenis Inferensi,  $F(3,231) = 26.29, p < .001$ , dengan inferensi MP yang paling sering didukung ( $M = 7.01, SD = 1.27$ ), diikuti oleh inferensi MT ( $M = 5.94, SD = 1.84$ ), inferensi AC ( $M = 5.82, SD = 2.10$ ), dan akhirnya inferensi DA ( $M = 4.76, SD = 2.54$ ). Tidak ada interaksi antara Jenis Inferensi dan Grup,  $F(3,231) < 1$ , menunjukkan kedua kelompok merespons tugas Inferensi Kondisional yang serupa pada Waktu 1.

Selanjutnya, perubahan tingkat dukungan setiap jenis inferensi dari waktu ke waktu dianalisis dengan ANOVA  $2 \times 4 \times 2$  dengan dua faktor dalam subjek: Waktu (awal dan akhir tahun) dan Jenis Inferensi (MP, DA, AC, MT), dan satu faktor antara mata pelajaran: Kelompok (matematika dan bahasa). Ini mengungkapkan interaksi tiga arah yang signifikan,  $F(3,228) = 7.48, p < .001, \eta^2 = .09$ , yang tetap signifikan setelah mengontrol Time 1 RAPM, Time 1 CRT dan pencapaian akademik sebelumnya,  $F(3,216) = 5.10, p = .002, \eta^2 = .07$  (lihat Gambar 5.7). Sarana dan standar deviasi untuk interaksi ini ditampilkan pada Tabel 5.4. Pada Waktu 2 mahasiswa matematika mendukung lebih banyak kesimpulan MP,  $t(42) = 2.42, p = .020, d = 0.41$ , dan DA lebih sedikit,  $t(42) = 3.98, p < .001, d = 0.67$ , AC,  $t(42) = 3.06, p = .004, d = 0.47$ , dan inferensi MT,  $t(42) = 2.88, p = .006, d = 0.45$  dibandingkan dengan Waktu 1. Sebaliknya, kelompok bahasa menunjukkan tidak ada perbedaan yang signifikan antara skor Waktu 1 dan Waktu 2 untuk setiap inferensi, meskipun ada sedikit peningkatan yang signifikan dalam jumlah kesimpulan DA yang didukung,  $t(34) = 1.80, p = 0.082, d = 0.31$ .



**Gambar 4.6:** Interaksi antara Kelompok dan Waktu pada nilai tes matematika (bar error menunjukkan  $\pm 1$  standar error dari mean).



**Gambar 4.7:** Rata-rata tingkat dukungan untuk masing-masing dari empat kesimpulan di setiap kelompok pada Waktu 1 dan Waktu 2 (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata).

Pada hakekatnya kelompok matematika menunjukkan peningkatan dukungan MP bersama dengan penurunan dukungan DA, AC dan MT, yang konsisten dengan interpretasi kondisional yang lebih cacat. Untuk menyelidiki ini secara formal, setiap indeks interpretasi dianalisis dengan ANOVA 2x2 dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan satu faktor antara mata pelajaran: Kelompok (matematika, Bahasa). Nilai indeks rata-rata untuk setiap kelompok pada setiap titik waktu ditunjukkan pada Gambar 4.8.

Untuk indeks kondisional material (MCI), ada interaksi yang signifikan antara Waktu dan Grup,  $F(1,76) = 11,86, p = .001, \eta^2 = .14$ , yang tetap signifikan ketika Waktu 1 RAPM, Waktu 1 CRT, dan skor pencapaian akademik sebelumnya dimasukkan sebagai kovariat,  $p = 0,007$ . Kelompok matematika menjadi lebih banyak materi,  $t(42) = 3.17, p = .003, d = 0.49$ , sedangkan kelompok bahasa tidak berubah,  $p = .092, d = 0.17$ .

Waktu dan Grup juga berinteraksi untuk indeks bikondisional (BCI),  $F(1,76) = 7.97, p = .006, \eta^2 = .10$ , meskipun ini hanya sedikit signifikan ketika kovariat dimasukkan,  $F(1,72) = 3,70, p = 0,058, \eta^2 = 0,05$ . Kelompok matematika menjadi kurang bikondisional,  $t(42) = 3.32, p = .002, d = 0.51$ , sedangkan kelompok bahasa tidak berubah,  $p = .500, d = 0,07$ .

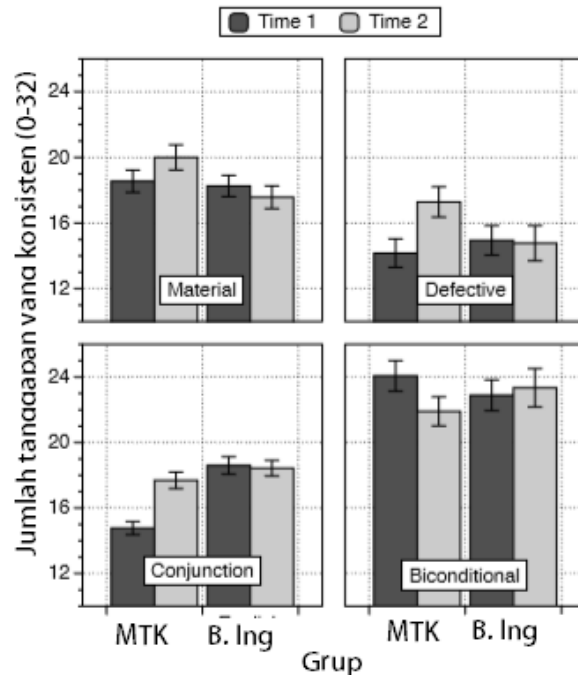
Untuk indeks kondisional cacat (DCI), Waktu dan Grup kembali berinteraksi,  $F(1,76) = 17,65, p < .001, \eta^2 = .19$ , dan ini tetap signifikan dengan kovariat,  $p = .002$ . Kelompok matematika menjadi lebih cacat,  $t(42) = 5.76, p < .001, d = 0.88$ , sedangkan kelompok bahasa tidak berubah,  $p = .767, d = 0.03$ .

Terakhir, untuk conjunctive conditional index (CCI), Waktu dan Grup juga berinteraksi,  $F(1,76) = 8.53, p = .005, \eta^2 = .10$ , yang tetap signifikan dengan kovariat,  $p = .014$ . Kelompok matematika menjadi lebih konjungtif,  $t(42) = 3.53, p = .001, d = 0.55$ , sedangkan kelompok bahasa tidak berubah,  $p = .69, d = 0.06$ .

Membandingkan ukuran efek dari analisis ini menegaskan bahwa perubahan dalam kelompok matematika paling baik dipahami sebagai peningkatan kecenderungan untuk mengadopsi cacat interpretasi dari kondisi ( $d = 0,88$  dibandingkan dengan  $d < 0,55$  untuk interpretasi lainnya). Seiring waktu kelompok matematika menjadi lebih mungkin untuk mendukung inferensi MP, tetapi kurang mungkin untuk mendukung inferensi DA, AC dan MT. Kelompok bahasa, di sisi lain, tidak mengubah salah satu indeks interpretasi.

**Tabel 4.4:** Jumlah rata-rata item yang didukung oleh jenis Inferensi, Kelompok dan Titik waktu dengan standar deviasi dalam tanda kurung.

Intervensi	Grup	Waktu 1	Waktu 2
MP	Matematika	7.12 (1.29)	7.60 (0.54)
	B. Inggris	6.91 (1.22)	7.06 (1.37)
DA	Matematika	5.00 (2.72)	3.86 (2.77)
	B. Inggris	4.49 (2.37)	5.06 (2.74)
AC	Matematika	5.76(2.22)	5.09(2.36)
	B. Inggris	5.83 (2.02)	5.83 (2.24)
MT	Matematika	6.19 (1.73)	5.36 (2.14)
	B. Inggris	5.66 (1.97)	5.40 (2.16)



**Gambar 4.8:** Rata-rata skor indeks interpretasi untuk setiap kelompok pada Waktu 1 dan Waktu 2 (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata).

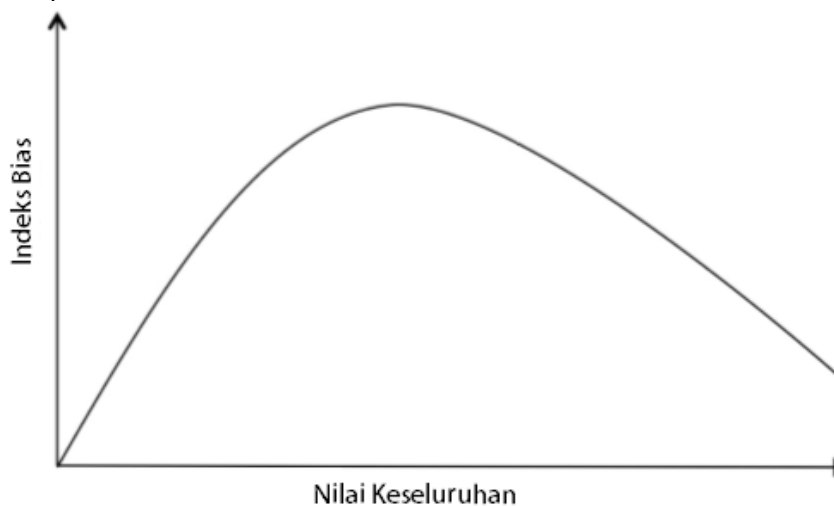
Skor NCI dan API. Skor NCI menjadi sasaran ANOVA  $2 \times 2$  dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan satu faktor antara mata pelajaran: Grup (matematika, bahasa dan tiga kovariat: pencapaian akademik sebelumnya, Waktu 1 RAPM skor dan skor Time 1 CRT. Ada sedikit interaksi yang signifikan antara Waktu dan Grup,  $F(1,72) = 3,60, p = 0,062, \eta^2 = 0,05$ . Menariknya, mahasiswa matematika menunjukkan NCI yang sedikit lebih tinggi pada Waktu 2 ( $M = 2.86, SD = 2.48$ ) dibandingkan pada Waktu 1 ( $M = 2.05, SD = 2.69$ ),  $t(43) = 1.81, p = .078, d = 0.31$ , sedangkan nilai mahasiswa bahasa tidak berubah (Waktu 1  $M = 2.09, SD = 3.08$ , Waktu 2  $M = 1.91, SD = 3.21$ ),  $t(35) = .33, p = .74, d = 0.06$ . Ini adalah kebalikan dari apa yang diharapkan mengingat temuan untuk skor keseluruhan - tampak bahwa selama tahun mahasiswa matematika menjadi lebih bias mendukung kesimpulan dengan kesimpulan negatif.

Skor API juga menjadi sasaran ANOVA  $2 \times 2$  dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan satu faktor antara mata pelajaran: Grup (matematika, bahasa), dan tiga kovariat: pencapaian akademik sebelumnya, Skor Waktu 1 RAPM dan skor Waktu 1 CRT. Hal ini tidak menunjukkan interaksi yang signifikan,  $F < 1$ .

Kompetensi penalaran dan bias. Dalam analisis di atas mempelajari matematika ditemukan terkait dengan peningkatan cacat dan interpretasi material dari kondisi dan peningkatan yang sedikit signifikan dalam NCI, dibandingkan dengan mempelajari bahasa. Ini agak berlawanan dengan intuisi. Sebuah interpretasi material atau cacat dapat dikatakan sebagai perbaikan pada interpretasi bikondisional (lihat diskusi di Bagian 2.3), dan orang akan berharap bahwa seiring dengan peningkatan kompetensi penalaran secara keseluruhan, kerentanan terhadap bias akan berkurang. Namun, bisa jadi bahwa tingkat pemahaman tertentu tentang tugas diperlukan sebelum seseorang dapat menunjukkan bias sistematis. Jika seorang penalaran memiliki sedikit pemahaman tentang logika kondisional, mereka mungkin merespons tugas secara tidak sistematis (menebak secara efektif) dalam hal ini tidak ada bias

sistematis yang dapat terjadi dan skor keseluruhan tidak akan jauh di atas tingkat peluang. Setelah pemahaman yang lebih canggih tetapi tidak sepenuhnya konsisten tentang kondisional tercapai, penalaran mungkin cukup sistematis untuk menunjukkan skor keseluruhan yang lebih tinggi dan mampu menunjukkan bias, tetapi tidak cukup kompeten untuk mengatasi bias setiap saat. Tentu saja, ketika tingkat pemahaman yang sangat tinggi tercapai, skor keseluruhan dan kemampuan untuk menghindari bias harus meningkat. Hubungan hipotetis ini ditunjukkan pada Gambar 4.9.

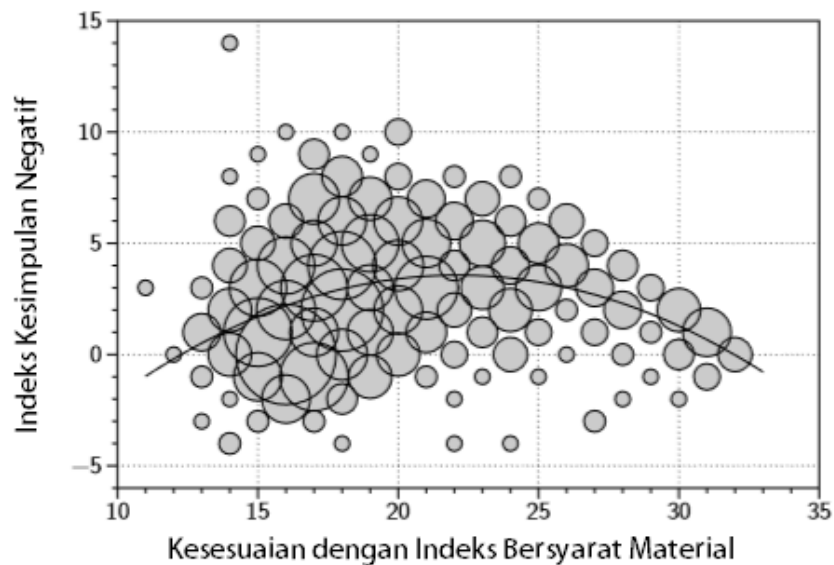
Untuk menguji hubungan yang diusulkan, data dari 656 peserta yang dikumpulkan dari lima studi terpisah (dibahas di atas) dianalisis. Data ini menjadi sasaran estimasi penyesuaian kurva, yang menegaskan bahwa kurva kuadrat memberikan kesesuaian yang lebih baik untuk hubungan antara MCI dan NCI pada tugas Inferensi Kondisional,  $R^2 = .09, F(2,653) = 33.61, p < .001$ , daripada hubungan linier,  $R^2 = .003, F(1,654) = 2.18, p = .14$ . Hubungan antara MCI dan NCI ditunjukkan pada Gambar 5.10. Hal ini juga berlaku untuk hubungan antara DCI dan NCI: kurva kuadrat memberikan kecocokan yang lebih baik untuk data,  $R^2 = .19, F(2,653) = 74.58, p < .001$ , daripada hubungan linier,  $R^2 = .03, F(1,654) = 18.78, p < .001$ . Hubungan antara DCI dan NCI ditunjukkan pada Gambar 4.11.



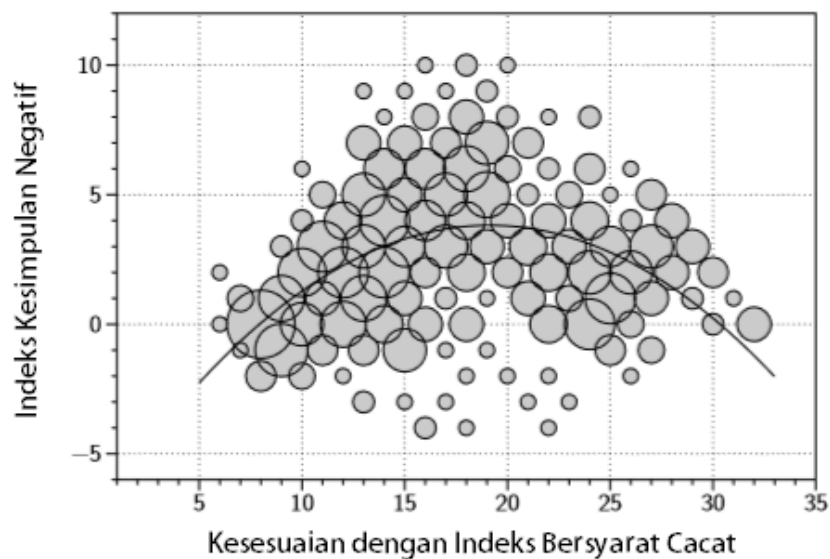
**Gambar 4.9:** Usulan hubungan antara skor keseluruhan dan indeks bias pada tugas penalaran berdasarkan temuan bahwa dari waktu ke waktu mahasiswa matematika menunjukkan skor total yang lebih tinggi dan indeks kesimpulan negatif sedikit lebih tinggi pada Tugas Inferensi Kondisional.

Namun, kasus NCI yang tinggi tidak dapat terjadi ketika skor indeks mendekati maksimum yang mungkin – jika sebagian besar inferensi dikategorikan secara konsisten maka tidak ada banyak ruang untuk pola yang dapat ditemukan dalam inferensi yang dikategorikan tidak konsisten. Mempertimbangkan batasan ini, cara lain untuk menguji hubungan yang diusulkan adalah dengan hanya melihat data yang skor indeksnya di bawah 75% dan ada ruang untuk terjadinya bias. Dalam rentang ini, hipopenelitiannya adalah bias akan meningkat seiring dengan skor indeks karena peserta menjadi cukup kompeten untuk bernalar secara sistematis tentang jawaban mereka, apakah itu benar atau salah. Untuk menguji hipopenelitian ini, skor MCI dikorelasikan dengan NCI untuk peserta yang skor MCI-nya kurang dari atau sama dengan 24 (dari 32). Ini mengungkapkan hubungan positif yang signifikan,  $r(554) = .25, p < .001$ , seperti yang diprediksi. Ada juga hubungan positif yang signifikan antara

DCI dan NCI untuk peserta yang DCI-nya kurang dari atau sama dengan 24,  $r(574) = 0,32, p < 0,001$ .



**Gambar 4.10:** Hubungan antara Indeks Kondisional Material dan Indeks Kesimpulan Negatif pada Tugas Inferensi Kondisional saat mengumpulkan data dari lima studi dengan garis yang paling cocok. Luas setiap lingkaran bertambah dengan jumlah orang di sel itu.

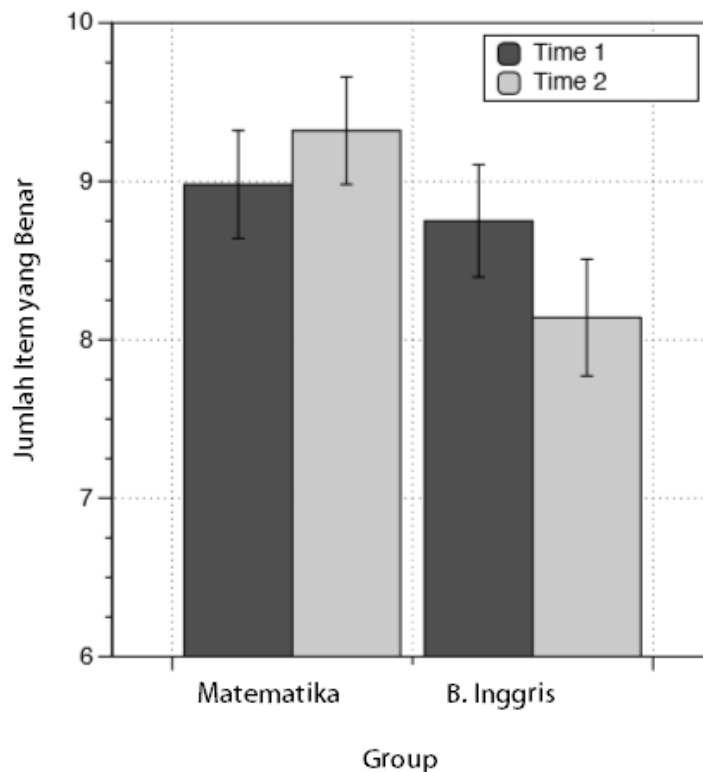


**Gambar 4.11:** Hubungan antara Indeks cacat kondisional dan indeks kesimpulan negatif pada tugas inferensi kondisional saat mengumpulkan data dari lima studi dengan garis yang paling cocok. Luas setiap lingkaran bertambah dengan jumlah orang di sel itu.

Estimasi penyesuaian kurva dan analisis korelasi konsisten dengan hipotesis bahwa ketika partisipan menjadi lebih berhasil dalam memikirkan deduksi dari pernyataan kondisional secara keseluruhan, mereka juga menjadi lebih rentan untuk menunjukkan bias sistematis dalam kesalahan mereka. Artinya, sampai mereka menjadi ahli dekat. Ini mungkin merupakan efek samping dari perpindahan dari respons acak yang efektif yang tidak memiliki alasan apa pun ke gaya penalaran yang lebih sistematis.

#### 4.7.4 Analisis silogisme.

Jumlah skor silogisme. Total skor silogisme menjadi sasaran ANOVA  $2 \times 2$  dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan satu faktor antara mata kuliah: Kelompok (matematika, bahasa), dan tiga kovariat: pencapaian akademik sebelumnya, Skor Waktu 1 RAPM dan skor Waktu 1 CRT. Hal ini menunjukkan interaksi yang signifikan,  $F(1,74) = 4,08, p = 0,047, \eta^2 = 0,05$ . Namun, nilai kelompok matematika tidak berubah antara Waktu 1 ( $M = 8,98, SD = 2,26$ ) dan Waktu 2 ( $M = 9,32, SD = 2,25$ ),  $t(44) = 1,06, p = .295, d = 0,15$ , dan nilai mahasiswa bahasa sedikit menurun antara Waktu 1 ( $M = 8,75, SD = 2,13$ ) dan Waktu 2 ( $M = 8,14, SD = 2,21$ ),  $t(36) = 1,90, p = 0,066, d = 0,28$ , lihat Gambar 4.12. Oleh karena itu interaksi tidak sejalan dengan TFD. Indeks Bias Keyakinan. Selanjutnya, analisis yang sama dilakukan dengan skor BBI. Tidak ada interaksi yang signifikan antara Waktu dan Grup,  $F(1,74) < 1$  (lihat Gambar 4.13).



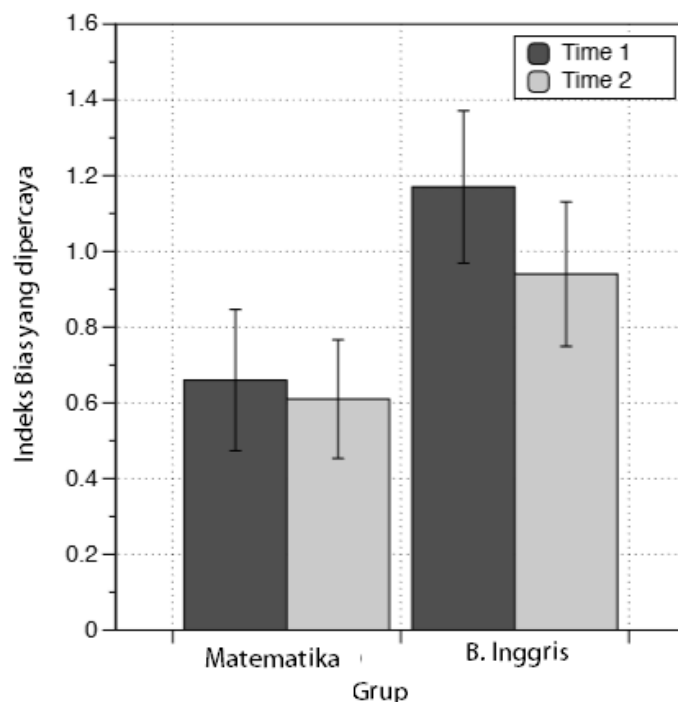
**Gambar 4.12:** Nilai rata-rata silogisme untuk setiap kelompok dan Waktu 1 dan Waktu 2 (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar dari rata-rata).

#### 4.7.5 Mekanisme perkembangan

Telah ditunjukkan bahwa kelompok matematika mengembangkan interpretasi kondisional yang lebih cacat dari waktu ke waktu. Namun, mereka tidak secara signifikan meningkatkan kemampuan penalaran silogistik atau menghindari bias keyakinan (walaupun mahasiswa bahasa mengalami penurunan kemampuan penalaran silogistik, menciptakan interaksi). Pada bagian ini, tingkat kognisi algoritmik dan reflektif Stanovich (2009a) dievaluasi sebagai mekanisme potensial dari perubahan interpretasi. Skor RAPM digunakan sebagai ukuran kecerdasan umum (tingkat algoritmik) dan skor intuitif terbalik pada CRT digunakan sebagai ukuran kecenderungan untuk menggunakan pemrosesan Tipe 2 (tingkat reflektif). Di bawah ini saya menyajikan model regresi yang memprediksi skor Time 2 DCI dari blok variabel berikut:

1. Pencapaian akademik sebelumnya, skor Waktu 1 DCI, skor Waktu 1 RAPM dan skor Waktu 1 CRT (semua kovariat);
2. Skor Change-in-RAPM dan skor *Change-in-CRT*, untuk menyelidiki apakah perubahan domain-umum bertanggung jawab atas skor Time 2 DCI di seluruh sampel;
3. Kelompok (matematika atau bahasa), untuk mengevaluasi apakah subjek yang dipelajari memprediksi kinerja Waktu 2 melebihi dan di atas perubahan domain-umum apa pun di seluruh sampel;
4. Dua istilah interaksi – Perubahan dalam Grup RAPM  $\times$  dan Perubahan dalam CRT  $\times$  Grup – untuk menyelidiki apakah perubahan domain-umum dalam grup matematika secara khusus bertanggung jawab atas skor Time 2 DCI.

Jika belajar matematika meningkatkan skor DCI dengan meningkatkan keterampilan pemrosesan domain-umum, maka istilah interaksi di blok terakhir harus menjelaskan jumlah varians yang signifikan – orang akan mengharapkan skor CRT dan RAPM berubah lebih banyak pada matematika, berbeda dengan kelompok bahasa, dan untuk memprediksi perkembangan DCI. Jika perubahan domain-general bertanggung jawab atas perubahan skor DCI secara independen dari subjek yang dipelajari, maka skor Waktu 1 atau Perubahan pada RAPM atau CRT harus menjadi prediktor yang signifikan. Jika studi matematika meningkatkan skor DCI melalui mekanisme selain kecerdasan atau disposisi bernalar, maka faktor Grup saja harus menjelaskan jumlah varians yang signifikan.



**Gambar 4.13:** Nilai indeks bias keyakinan rata-rata untuk setiap kelompok dan Waktu 1 dan Waktu 2 (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata).

Model regresi hierarkis disajikan pada Tabel 4.5. Variabel kontrol di blok pertama prediktor menyumbang 65,6% dari varians dalam skor Time 2 DCI,  $p < 0,001$ . Skor Change-in-RAPM dan Change-in-CRT ditambahkan pada blok kedua dan menyumbang 3,2% tambahan yang signifikan dari varians,  $p = 0,036$ . Kelompok (matematika = 1, Bahasa = 0) dimasukkan di blok ketiga dan menyumbang 2,1% tambahan dari varians,  $p = 0,032$ . Menambahkan dua



istilah interaksi di blok terakhir menyumbang kurang dari 0,5% varians tambahan,  $p = 0,581$ . Dalam model akhir, satu-satunya prediktor yang signifikan adalah Waktu 1 DCI ( $p < 0,001$ ) dan Grup ( $p = 0,024$ ).

Model regresi tidak mendukung hipotesis bahwa mempelajari matematika meningkatkan kesesuaian cacat interpretasi dari kondisi faktor-faktor umum dalam domain kecerdasan (pada tingkat algoritmik) atau disposisi bernalar (pada tingkat reflektif). Sebaliknya, mekanisme perbaikan mungkin domain-spesifik 'mindware', faktor tingkat heuristik, atau faktor pada tingkat kognisi algoritmik atau reflektif yang tidak termasuk di sini, seperti fungsi eksekutif. Variabel kontrol yang termasuk dalam Blok 1, seperti yang diharapkan, merupakan prediktor yang signifikan. Namun, bukan kasus matematika yang mempengaruhi kecerdasan atau disposisi bernalar dengan cara yang diprediksi Waktu 2 DCI: istilah interaksi di blok terakhir menjelaskan varians hampir nol.

**Tabel 5.5:** Analisis regresi hierarkis yang memprediksi skor Indeks Kondisional Waktu 2 yang Rusak.  $^+p < .1$ ,  $*p < .05$ ,  $**p < .001$ .

Model	R <sup>2</sup>	$\Delta R^2$	Prediktor	$\beta$
1	.66	.66**	DCI Waktu 1 Raven Waktu 1 CRT waktu 1 Pencapaian akademik sebelumnya	.71** 03 .22** .01
2	.69	.03*	DCI Waktu 1 Raven Waktu 1 CRT waktu 1 Pencapaian akademik sebelumnya Perubahan-dalam-RAPM Perubahan-dalam-CRT	.72** .13 .18 .01 .20* .06
3	.71	.02*	DCI Waktu 1 Raven Waktu 1 CRT waktu 1 Pencapaian akademik sebelumnya Perubahan-dalam-RAPM Perubahan-dalam-CRT Kelompok (matematika = 1, Bahasa = 0)	.76** .06 .13 -.01 .15 <sup>+</sup> .04 .17*
4	.71	.01	DCI Waktu 1 Raven Waktu 1 CRT waktu 1 Pencapaian akademik sebelumnya Perubahan-dalam-RAPM Perubahan-dalam-CRT Kelompok (matematika = 1, Bahasa = 0) Perubahan RAPM $\times$ Grup Perubahan CRT $\times$ Grup	.75** .07 .12 -.01 .14 .09 .20* .02 -.09

#### 4.7.6 Pengaruh mata pelajaran ilmu lainnya

Pada tingkat AS, mahasiswa cenderung mempelajari empat mata kuliah, dan mereka yang belajar matematika seringkali juga mempelajari mata kuliah IPA. Di Inggris pada tahun 2009, misalnya, 63% mahasiswa yang belajar matematika pasca-wajib juga mempelajari satu atau lebih kimia, biologi, dan fisika (Royal Society, 2011). Untuk menyelidiki apakah matematika itu unik sebagai disiplin formal yang memprediksi perkembangan dalam inferensi kondisional, maka jumlah mata pelajaran ilmu inti yang dipelajari setiap peserta (fisika, kimia, biologi) dikorelasikan dengan perubahan skor DCI dalam kelompok matematika. Hal ini mengungkapkan hubungan yang tidak signifikan,  $r(43) \text{ Spearman} = .16, p = .316$ , yang tetap tidak signifikan setelah mengontrol RAPM Waktu 1 peserta, skor CRT, dan skor pencapaian akademik sebelumnya,  $pr(37) = .22, p = .188$ . Dalam kumpulan data saat ini tidak ada bukti bahwa mempelajari mata kuliah ilmu non-matematika berbeda dengan ilmu matematika dalam perubahan DCI.

#### 4.8 DISKUSI

Bab ini bertujuan untuk menjelaskan dua pertanyaan yang penting untuk penelitian ini dan TFD: (a) apakah belajar matematika pada tingkat lanjutan terkait dengan peningkatan keterampilan penalaran? dan (b) jika ada peningkatan seperti itu, bagaimana mekanismenya? Pertanyaan-pertanyaan ini dijawab dengan studi longitudinal yang mengikuti perkembangan kemampuan penalaran kondisional dan silogistik pada mahasiswa matematika dan bahasa tingkat AS. Hasilnya menunjukkan bahwa (a) mahasiswa matematika berubah dalam perilaku penalaran kondisional ke tingkat yang lebih besar daripada mahasiswa bahasa, (b) perubahan yang paling nyata ditandai dengan mahasiswa matematika menjadi cacat interpretasi mereka terhadap kondisional, dan (c) bahwa mekanisme pengembangannya tampaknya tidak bersifat umum.

#### 4.9 PENGEMBANGAN KETRAMPILAN PENALARAN

Ditemukan di sini bahwa perilaku penalaran kondisional mahasiswa matematika menjadi lebih sejalan dengan interpretasi kondisional yang rusak dari waktu ke waktu, sedangkan perilaku penalaran mahasiswa bahasa tidak berubah. Inglis dan Simpson (2009a) menemukan bahwa, dibandingkan dengan mahasiswa matematika kecerdasan mahasiswa matematika lebih cerdas secara normatif pada tugas inferensi kondisional, tetapi mereka tidak berubah selama satu tahun studi matematika. Para penulis menyarankan bahwa perbedaan awal mungkin disebabkan oleh salah satu dari tiga kemungkinan: studi matematika pasca-wajib tetapi pra-universitas mengembangkan keterampilan penalaran; menyaring lebih banyak alasan material ke dalam studi matematika; atau perbedaan antarkelompok yang tidak terkait dengan kecerdasan, seperti dalam disposisi bernalar. Temuan ini konsisten dengan kemungkinan pertama, bahwa studi matematika pasca-wajib tetapi pada saat pra-universitas dikembangkan keterampilan penalaran kondisional. Pada awal pendidikan pasca wajib, mahasiswa yang belajar matematika tidak berbeda dari mahasiswa non-matematika pada tugas inferensi kondisional, tetapi mereka menjadi berubah setelah satu tahun belajar. Perubahan ini bukan karena perbedaan antar kelompok dalam disposisi bernalar atau skor kecerdasan awal yang berubah.

Penalaran mahasiswa matematika tidak sepenuhnya menjadi lebih normatif seperti yang kita harapkan, atau mempunyai pemahaman logika yang lebih baik. Ada peningkatan

penolakan inferensi DA, AC dan MT dan peningkatan penerimaan inferensi MP, yang mencerminkan langkah menuju apa yang disebut 'cacat' interpretasi dari kondisi. Ini tidak sepenuhnya mengejutkan mengingat sifat matematika: Houston (2009) berpendapat bahwa sebagian besar pernyataan matematika berbentuk 'jika pernyataan A benar, maka pernyataan B benar', bahkan jika pernyataan tersebut sangat disamakan (hal. 63). Dia juga berpendapat bahwa dalam matematika A dianggap benar, bahkan jika itu jelas tidak benar, dan kebenaran atau kesalahan B kemudian disimpulkan. Karena kurikulum tingkat AS tidak menyertakan referensi eksplisit ke logika kondisional atau kondisional materi, masuk akal bahwa paparan pernyataan 'jika kemudian' implisit, di mana anteseden diasumsikan benar, dapat menyebabkan interpretasi yang rusak dari kondisi, dimana kasus anteseden palsu dianggap tidak relevan. Sejalan dengan hipopenelitian ini, Hoyles dan Kuchemann (2002) berpendapat bahwa cacat interpretasi sebenarnya lebih tepat untuk kelas matematika daripada interpretasi materi untuk alasan yang sama yang diajukan oleh Houston. Lebih lanjut, Inglis dan Simpson (2009b) menemukan bahwa sekelompok mahasiswa sarjana matematika, yang tentu saja sangat berhasil pada matematika tingkat A, cenderung memiliki cacat interpretasi kondisional daripada materi lain.

Ada hal yang mengejutkan bahwa belajar matematika dikaitkan dengan peningkatan keterampilan penalaran kondisional, sebab Cheng et al. (1986) tidak menemukan perbaikan dalam penalaran kondisional bahkan setelah peserta mereka mempelajari kuliah logika formal. Salah satu penjelasan yang mungkin untuk perbedaan ini adalah ukuran yang digunakan oleh Cheng et al. (1986) tidak cocok untuk mendeteksi perbaikan. Mereka menggunakan empat Tugas Seleksi Wason, tiga di antaranya dikontekstualisasikan, untuk mengukur kemampuan penalaran kondisional. Sejak itu telah disarankan bahwa Tugas Seleksi Wason yang dikontekstualisasikan, mungkin tidak mengukur penalaran kondisional sama sekali (Sperber et al., 1995, 2002) dan jadi ada kemungkinan bahwa Cheng et al. (1986) sama sekali tidak mengukur peningkatan yang benar-benar terjadi.

Mungkin juga jika Cheng et al. (1986) telah melihat secara khusus pada indeks kondisional yang rusak mereka mungkin telah menemukan perubahan - dalam studi saat ini ada pergeseran menuju interpretasi material, tapi lebih kecil dari pergeseran menuju interpretasi yang rusak. Jika peserta Cheng et al menjadi lebih cacat dari waktu ke waktu, itu tidak akan tercermin oleh kinerja pada Tugas Seleksi – itu akan mendorong peserta untuk memilih hanya kartu anteseden yang benar dan tidak ada yang lain, yang akan dianggap salah dalam analisis Cheng et al. , bersama dengan interpretasi bikondisional. Kebetulan, ini adalah pola respons yang ditemukan dalam studi Inglis dan Simpson (2004) tentang mahasiswa matematika sarjana. Dibandingkan dengan kelompok kontrol, mahasiswa matematika Inglis dan Simpson (2004) memberikan lebih banyak jawaban true antecedent (p) only dan lebih sedikit true antecedent dan true consequent (p dan q) jawaban.

#### **4.10 KOMPETENSI DAN BIAS DALAM TUGAS INFERENSI KONDISIONAL**

Analisis tambahan menggunakan sekumpulan besar data yang dikumpulkan dari beberapa penelitian menguji hipopenelitian bahwa indeks kesimpulan negatif meningkat seiring dengan peningkatan materi dan interpretasi yang salah, hingga titik konsistensi yang hampir lengkap dalam jenis inferensi. Untuk mendukung hal ini, hubungan antara MCI dan NCI dan antara DCI dan NCI lebih sesuai dengan kurva kuadratik daripada fungsi linier. Hubungan

ini menunjukkan bahwa ketika orang menjadi lebih sistematis dalam penalaran mereka, mereka juga menjadi lebih rentan terhadap bias umum.

Ketika orang merespons secara tidak sistematis, baik karena mereka tidak memahami kondisi atau karena mereka tidak terlibat dengan tugas, tidak mungkin bagi mereka untuk menunjukkan bias kesimpulan negatif karena memerlukan beberapa pemahaman dan konsistensi dalam menanggapi deduksi serupa. Orang-orang yang memahami kondisi dan merespons dengan beberapa konsistensi terhadap deduksi serupa setidaknya mampu menunjukkan bias kesimpulan negatif, dan mungkin atau mungkin tidak dapat mengatasinya. Akhirnya, mereka yang memiliki pemahaman yang konsisten tentang kondisi mampu mengatasi bias kesimpulan negatif dan menjawab semua atau sebagian besar item sesuai dengan interpretasi utama mereka. Singkatnya, pola ini mungkin mencerminkan lintasan perkembangan dari tidak sistematis menanggapi lebih sistematis tetapi bias menanggapi penalaran sepenuhnya sistematis.

#### **4.11 KETERBATASAN**

Mahasiswa yang belajar matematika dalam sampel juga cenderung mempelajari mata pelajaran IPA inti lainnya (fisika, biologi, kimia). Tidak mungkin untuk memisahkan hubungan potensial antara perubahan perilaku penalaran kondisional dan studi matematika, fisika, kimia dan biologi dalam sampel saya, karena sejumlah kecil mahasiswa yang belajar hanya satu mata kuliah. Namun, tidak ada bukti bahwa semakin banyak mata pelajaran ilmu yang dipelajari seseorang, semakin banyak perilaku penalaran mereka yang berubah, jadi meskipun pembaur ini tidak ideal tampaknya tidak menjadi faktor pembatas yang parah. Untuk menyelidiki potensi masalah pembaur, ukuran sampel yang jauh lebih besar akan diperlukan untuk memastikan bahwa cukup banyak peserta yang mempelajari satu mata pelajaran ilmu saja. Namun, penelitian yang dilaporkan dalam Bab 6 menyelidiki peran belajar matematika tanpa mata pelajaran ilmu pada mahasiswa sarjana dan menyarankan bahwa hubungan yang ditemukan di sini masih berlaku.

Sebuah kelemahan substansial untuk desain penelitian ini adalah untuk eksperimen semu yang bertentangan dengan eksperimen yang sebenarnya. Peserta memilih mata pelajaran tingkat AS mereka sebelum penelitian dimulai – tidak praktis atau etis untuk secara acak menugaskan mahasiswa untuk belajar matematika atau bahasa. Meskipun tidak ada perbedaan antara kelompok dalam perilaku penalaran pada Time 1, desain eksperimen semu berarti bahwa tidak mungkin untuk membangun hubungan sebab akibat antara subjek yang dipelajari dan peningkatan keterampilan penalaran.

Seperti yang dibahas dalam Bab 3, penugasan non-acak ke kondisi berarti bahwa tidak semua variabel pengganggu dapat dicegah untuk memengaruhi hubungan yang dipelajari. Kecerdasan dan disposisi bernalar, dua kemungkinan pembaur dalam hubungan antara subjek yang dipelajari dan keterampilan penalaran, diukur dan dikontrol secara statistik, tetapi kontrol statistik tidak seefektif penugasan acak untuk kondisi (Christensen, 2000) dan, yang lebih penting, mungkin ada menjadi variabel pengganggu lain yang tidak dipertimbangkan. Cara yang paling efektif untuk mempelajari sebab-akibat di TFD adalah dengan secara acak menugaskan peserta ke kursus matematika atau mata pelajaran non-matematis dan melacak perkembangan mereka dalam penalaran selama periode waktu yang lama. Namun, ini jelas tidak praktis atau etis dalam penilaian taruhan tinggi.

Dalam kaitannya dengan kebijakan pendidikan, desain eksperimen semu dari penelitian ini juga berarti bahwa hasilnya mungkin tidak berlaku untuk kurikulum di mana wajib belajar matematika sampai usia 18 tahun. Para peserta dalam penelitian ini telah memilih untuk belajar matematika dan sehingga mereka mungkin menikmatinya dan terlibat dengan kuliah. Di mana mahasiswa diminta untuk belajar matematika sampai usia 18 tahun ini mungkin tidak terjadi dan mungkin hanya mereka yang menikmati dan terlibat dengan kursus melihat manfaat untuk penalaran yang ditemukan. Arah yang berguna untuk penelitian masa depan adalah membandingkan perkembangan penalaran mahasiswa yang mempelajari matematika dalam kurikulum yang wajib dan tidak wajib.

#### **4.12 REVISI STATUS TEORI DISIPLIN FORMAL**

Teori Disiplin Formal menunjukkan bahwa mempelajari matematika meningkatkan kemampuan seseorang untuk bernalar. Meskipun TFD membuat klaim besar tentang hubungan antara matematika dan penalaran, TFD tidak membahas tentang mekanisme hubungan yang mungkin terjadi. Sampai saat ini, TFD telah dianggap benar dengan pengawasan minimal. Di sini, beberapa bukti disajikan yang sebagian konsisten dengan TFD, dan lebih jauh lagi, beberapa mekanisme potensial untuk peningkatan keterampilan penalaran diselidiki.

Dalam studi yang disajikan di sini ditemukan bahwa perilaku penalaran kondisional mahasiswa matematika berubah ke tingkat yang lebih besar daripada mahasiswa nonmatematika, tetapi mahasiswa matematika tidak menunjukkan perubahan dalam penalaran silogistik mereka atau kerentanan terhadap bias keyakinan. Saya juga tidak menemukan bukti bahwa mekanisme perubahan dalam keterampilan penalaran kondisional adalah perubahan ke tingkat kognisi algoritmik atau reflektif, seperti yang ditentukan dalam model Stanovich (2009a). Di sini, RAPM adalah satu-satunya ukuran pemrosesan algoritmik, tetapi fungsi eksekutif adalah aspek yang dapat dipisahkan dari tingkat algoritmik. Fungsi eksekutif mengacu pada efisiensi dalam melakukan pemrosesan informasi secara sadar – mis. memperbarui informasi dalam memori kerja, beralih di antara tugas-tugas, dan menghambat informasi atau tanggapan yang tidak relevan, dan telah terbukti terpisah dari kecerdasan cairan. Hubungan antara fungsi eksekutif dan perilaku penalaran kondisional diselidiki dalam Bab 8.

Kemungkinan lain adalah respon mahasiswa matematika yang cacat berasal dari tingkat kognisi heuristik. Hal ini diselidiki dalam Bab 7, di mana perilaku penalaran kondisional mahasiswa matematika dan non-matematika diukur di bawah batas waktu yang ketat serta di bawah batas waktu jadi lihat bagaimana perubahan perilaku saat pemrosesan dibatasi pada tingkat heuristik. Bab 7 menyajikan bukti bahwa kemampuan mahasiswa matematika dengan penalaran kondisional tergantung pada bentuk linguistik dari pernyataan kondisional, mendukung hipotesis bahwa paparan pernyataan 'jika kemudian' bertanggung jawab atas perubahan yang bertentangan dengan pemahaman umum logika kondisional. Selanjutnya, bagaimanapun, sebuah studi longitudinal menyelidiki perubahan dalam perilaku penalaran kondisional pada mahasiswa sarjana disajikan.

#### **4.13 RINGKASAN DAN PENEMUAN BARU**

1. Perilaku penalaran kondisional mahasiswa matematika lebih sesuai dengan cacat interpretasi dan kurang sesuai dengan interpretasi bikondisional dari pernyataan

kondisional setelah satu tahun studi di tingkat AS, dibandingkan dengan mahasiswa bahasa.

2. Tidak ada bukti bahwa perilaku penalaran silogistik mahasiswa matematika atau kerentanan terhadap bias keyakinan berubah dari waktu ke waktu.
3. Mekanisme perubahan perilaku penalaran kondisional mahasiswa matematika tampaknya tidak menjadi dasar atau skor yang berubah pada ukuran kecerdasan atau disposisi bernalar.
4. Perubahan terbesar dalam perilaku penalaran kondisional mahasiswa matematika adalah dalam bentuk adopsi yang lebih besar dari kondisi cacat yang mungkin mencerminkan praktik dengan pernyataan implisit 'jika maka' dalam matematika di mana mahasiswa diharapkan untuk mengasumsikan  $p$  dan alasan tentang  $q$ .

## BAB 5

### MENGEMBANGKAN KEMAMPUAN BERNALAR PASCASARJANA

#### 5.1 MENGUJI TEORI DISIPLIN FORMAL

Pada bab 4 menyajikan studi yang menyelidiki perkembangan keterampilan penalaran dalam matematika tingkat AS dan mahasiswa bahasa. Ditemukan bahwa mahasiswa matematika menjadi semakin cacat dalam perilaku penalaran mereka, yaitu mereka menjadi lebih mungkin untuk menolak inferensi DA, AC dan MT, dan lebih mungkin untuk menerima inferensi MP. Ditemukan juga bahwa mahasiswa matematika tidak meningkat dalam penalaran dengan silogisme tematik. Ketidaksiharian antara peningkatan penalaran kondisional abstrak dan kurangnya peningkatan silogisme tematik dapat disebabkan oleh aspek konteks/abstrak atau aspek kondisional/silogisme. Ini akan diklarifikasi dalam penelitian yang dilaporkan di bawah ini.

Tujuan dari bab ini adalah untuk menyelidiki perkembangan keterampilan penalaran pada mahasiswa pascasarjana matematika. Karena studi dan motivasinya sangat mirip dengan studi AS, banyak latar belakang dan deskripsi materi yang relevan telah muncul pada bab 4. Untuk menghindari pengulangan, hanya gambaran singkat tentang latar belakang yang relevan dan pembenaran studi yang diberikan di bawah ini.

Terlalu sedikit penelitian terdahulu yang menyelidiki perkembangan keterampilan penalaran dalam kaitannya dengan studi matematika semuanya terfokus pada mahasiswa sarjana. Lehman dan Nisbett (1990) menguji mahasiswa sarjana level AS pada berbagai jenis penalaran pada awal dan akhir empat tahun studi mereka dan menemukan bahwa mahasiswa ilmu alam (yang mengambil modul matematika paling banyak) bernalar lebih sesuai dengan materi kondisional di akhir gelar mereka. Selanjutnya, jumlah modul matematika yang diambil dikorelasikan dengan tingkat perubahan.

Inglis dan Simpson (2008) menemukan bahwa saat masuk ke universitas, pascasarjana matematika beralasan lebih normatif daripada kelompok pembandingan pada tugas Inferensi Kondisional. Namun, dalam studi lanjutan Inglis dan Simpson (2009a) menemukan bahwa meskipun mahasiswa matematika kembali mengungguli perbandingan sarjana saat masuk ke universitas, alasan mereka tidak berubah selama satu tahun studi. Para penulis menyarankan bahwa belajar matematika di tingkat A mungkin telah menyebabkan perbedaan awal antara kelompok, dan ini didukung oleh penelitian yang disajikan dalam bab 4 dari buku ini. Namun, penulis tidak menyelidiki perbedaan atau perubahan dalam interpretasi kondisional yang cacat, sehingga ada kemungkinan ada perubahan yang tidak terdeteksi. Namun demikian, ada perbedaan antara temuan Lehman dan Nisbett (1990) dan Inglis dan Simpson (2009a) untuk interpretasi materi kondisional, dan pada bab ini penulis akan menambahkan bukti yang dapat memperjelas perbedaan tersebut. Selain berkontribusi pada basis bukti terbatas yang berkaitan dengan pengembangan keterampilan penalaran pada mahasiswa matematika sarjana, studi yang disajikan di bawah ini akan membahas dua masalah yang muncul dalam studi AS yang disajikan dalam bab 4.

Pertama, dalam sampel AS, ilmu dan matematika dikacaukan sehingga tidak mungkin untuk mengisolasi pengaruh pembelajaran matematika terhadap penalaran. Di Indonesia, mahasiswa cenderung mempelajari hanya satu mata kuliah pada tingkat S2, sehingga studi

yang disajikan di sini memungkinkan efek matematika diselidiki secara terpisah. Masalah ini juga dapat menjadi sumber perbedaan yang dijelaskan di atas – Lehman dan Nisbett (1990) menemukan bukti mereka untuk hubungan antara matematika dan penalaran kondisional pada mahasiswa tingkat AS, di mana pada umumnya mahasiswa S1 untuk mempelajari mata kuliah yang berbeda sebagai pelengkap dari program studi mereka, meskipun derajatnya kecil. Tidak ada bukti hubungan antara mempelajari mata kuliah matematika dengan perubahan DCI pada mahasiswa tingkat AS, tetapi hal itu disebabkan banyaknya campuran mata kuliah yang dipelajari oleh para mahasiswa. Lehman dan Nisbett (1990) menyatakan bahwa campuran itu berpotensi menjadi alasan untuk menemukan tidak adanya perubahan seperti hasil temuan Inglis dan Simpson (2009a).

Kedua, pada pembelajaran AS matematika mahasiswa hanya berubah pada penalaran abstrak kondisional, tidak pada penalaran silogistik tematik. Kurangnya perbaikan silogisme bisa jadi karena penggunaan konteks – mungkin belajar matematika hanya memberikan keuntungan untuk bernalar tentang masalah abstrak – atau bisa juga karena silogisme – mungkin belajar matematika hanya memberikan keuntungan untuk memikirkan pernyataan kondisional (peningkatan interpretasi yang rusak dari kondisional mendukung penjelasan terakhir). Masalah ini akan diklarifikasi dalam studi yang disajikan di sini. Alih-alih menyelesaikan tugas silogisme tematik, peserta akan menyelesaikan tugas inferensi kondisional tematik yang bentuknya sangat mirip dengan tugas penalaran kondisional abstrak. Jika mahasiswa matematika mengubah versi abstrak tetapi bukan versi tematik, itu akan menunjukkan bahwa konteks mengganggu penalaran mereka. Jika mereka berubah pada versi abstrak dan tematik (khususnya, jika penalaran mereka menjadi lebih cacat pada keduanya), itu menunjukkan bahwa belajar matematika hanya mengubah interpretasi 'jika' dan itulah sebabnya tidak ada perubahan pada tugas Silogisme di mahasiswa tingkat AS. Namun, ada kemungkinan alternatif dalam kasus terakhir adalah penalaran dalam konteks keterampilan yang muncul di kemudian hari dalam studi matematika daripada penalaran abstrak - pada titik di antara tingkat AS dan tahun pertama mendapat gelar sarjana.

Tujuan lebih lanjut dari studi ini, seperti studi AS, adalah untuk mengidentifikasi mekanisme potensial untuk setiap perbaikan yang terjadi. Seperti sebelumnya, ukuran kecerdasan pada tingkat algoritmik dan disposisi bernalar reflektif pada tingkat kognisi reflektif dalam model Stanovich (2009a) akan dimasukkan untuk tujuan ini.

## 5.2 RINGKASAN

Singkatnya, ada tiga pertanyaan penelitian yang akan dibahas dalam bab ini: (a) mempelajari matematika dan ilmu sosial di tingkat sarjana hanya terkait dengan perubahan keterampilan penalaran kondisional abstrak, (b) mempelajari matematika dan ilmu sosial di tingkat sarjana terkait dengan perubahan dalam keterampilan penalaran kondisional tematik dan (c) dapatkah peningkatan apa pun yang ditemukan dikaitkan dengan perubahan kecerdasan atau disposisi bernalar reflektif? Berdasarkan temuan dari studi tingkat AS, dapat dihipotesiskan bahwa:

1. Belajar matematika di tingkat sarjana akan dikaitkan dengan lebih banyak cacat, dan pada tingkat yang lebih rendah, lebih banyak penalaran kondisional abstrak material,
2. Perubahan dalam penalaran kondisional abstrak akan diprediksi oleh faktor Kelompok di atas kecerdasan dan/atau disposisi bernalar.



Tidak jelas apakah mahasiswa matematika akan meningkat dalam penalaran kondisional tematik, tetapi jika mereka melakukannya, hipopenelitian yang sama kemungkinan akan berlaku untuk peningkatan tersebut.

### **5.3 METODE**

#### **5.3.1 Desain**

Mahasiswa sarjana matematika dan psikologi mengambil bagian pada awal dan akhir tahun pertama studi mereka. Peserta telah memilih sendiri program gelar dan studi mengambil desain kuasi-eksperimental. Serangkaian tugas yang sama diberikan pada kedua titik waktu untuk memungkinkan penyelidikan perkembangan yang membujur.

#### **5.3.2 Partisipan**

Delapan puluh tiga mahasiswa matematika dan 64 mahasiswa psikologi mengikuti Time 1. Kelompok matematika ini hanya terdiri dari mahasiswa S1 yang sedang menempuh mata kuliah matematika unggulan empat tahun ( $N = 66$ ) dan program pascasarjana matematika dua tahun (MMath,  $N = 17$ ). Semua peserta adalah mahasiswa tahun pertama di Universitas terkemuka dan mengambil bagian secara sukarela tanpa dibayar.

#### **5.3.3 Silabus Mata Kuliah Matematika**

Pada tahun pertama gelar kehormatan tunggal matematika dan MMath mahasiswa mengambil modul wajib berikut: Kalkulus, Aljabar linier, Geometri, Vektor dan bilangan kompleks, Penalaran matematika, Pengantar matematika terapan, Aplikasi komputer dalam matematika, Barisan dan deret, Persamaan diferensial dan Probabilitas dan statistik pengantar. Bertentangan dengan silabus tingkat A yang dibahas dalam Bab 4, modul sarjana 'Penalaran matematis' mencakup berbagai aspek logika, termasuk pernyataan kondisional dan tabel kebenaran.

### **5.4 PENGUKURAN**

#### **5.4.1 Inferensi Kondisional Abstrak**

Peserta menyelesaikan tugas Inferensi Kondisional yang sama (Evans et al., 1995) yang digunakan dalam bab 4, terdiri dari 32 item abstrak dari empat jenis inferensi: modus ponens (MP), penolakan anteseden (DA), penegasan konsekuensi (AC) dan modus tollens (MT). Setengah dari item menggunakan negasi eksplisit (misalnya "bukan 5") dan setengahnya menggunakan negasi implisit (misalnya "bukan 5" direpresentasikan sebagai, misalnya, 6). Empat indeks interpretasi diambil: MCI, DCI, BCI dan CCI. Instruksi yang diberikan identik dengan yang digunakan oleh Evans et al. (1995). Contoh item ditunjukkan pada Gambar 6.1.

#### **5.4.2 Inferensi Kondisional Tematik**

Versi tematik tugas inferensi kondisional dibuat berdasarkan tugas Evans, Handley, Neilens dan Over (2010). Partisipan memutuskan apakah kesimpulan harus diikuti dari aturan dan premis dalam 16 item. Ada empat item MP, empat MT, empat DA dan empat AC. Dua dari setiap jenis inferensi disajikan dalam konteks yang dapat dipercaya dan dua dalam konteks yang tidak dapat dipercaya. Contoh item MT yang sulit dipercaya ditunjukkan pada Gambar 5.2.

Jika hurufnya S maka angkanya adalah 6 Nomornya bukan 6 Kesimpulan: Hurufnya bukan S	Jika hurufnya M maka angkanya 4 Surat itu bukan M Kesimpulan: Angkanya bukan 4
<input type="radio"/> Ya <input type="radio"/> Tidak	<input type="radio"/> Ya <input type="radio"/> Tidak
a) Modus tollens	b) Penolakan dari pendahulunya

**Gambar 5.1:** Contoh item dari tugas Inferensi Kondisional Abstrak.

Semua negasi diwakili secara eksplisit: kurangnya negasi implisit adalah alasan mengapa ada 16 item, bukannya 32 item. Langkah-langkah yang dilakukan adalah MCI, DCI, BCI, CCI, dan indeks bias keyakinan (BBI), seperti pada tugas silogisme pada Bab 5. BBI dihitung sebagai jumlah item yang konsisten diterima (dapat dipercaya/valid, sulit dipercaya/tidak valid) dikurangi jumlah item tidak konsisten yang diterima (dapat dipercaya/tidak valid, tidak dapat dipercaya/valid). Rentang yang mungkin untuk skor BBI adalah -8 hingga +8, dengan skor positif menunjukkan tingkat bias keyakinan (lebih dibujuk oleh keyakinan daripada validitas dalam item yang tidak konsisten).

#### **5.4.3 Matriks Progresif Tingkat Lanjut Raven**

Sebuah subset 18 item dari Raven's Advanced Progressive Matrices (RAPM) dengan batas waktu 15 menit (S'a et al., 1999) digunakan sebagai ukuran kecerdasan umum (pada tingkat pemrosesan algoritmik, Stanovich, 2009a, lihat Lampiran SEBUAH).

Asumsikan hal berikut ini benar: Jika utang dunia ketiga dibatalkan maka kemiskinan akan semakin parah. Mengingat bahwa premis berikut ini juga benar: Kemiskinan dunia ketiga tidak memburuk. Apakah perlu bahwa: Utang dunia ketiga tidak dibatalkan. <input type="radio"/> Ya
--

**Gambar 5.2:** Contoh item dari tugas Inferensi Kondisional Tematik

#### **5.4.4 Tes Refleksi Kognitif**

Jumlah tanggapan intuitif yang diberikan kepada tiga item CRT (Frederick, 2005) digunakan sebagai ukuran kinerja kecenderungan untuk menggunakan pemrosesan Tipe 2 (pada tingkat reflektif, Toplak et al., 2011; Stanovich, 2009a). Skor dibalik sehingga skor yang lebih tinggi mewakili kinerja yang lebih normatif, sejalan dengan ukuran lainnya. Pertanyaan-pertanyaan tersebut dicampur secara acak dengan tiga soal matematika sederhana dengan panjang yang sama dari subtes Masalah Terapan Woodcock-Johnson III seperti yang dijelaskan dalam Bab 4.

#### **5.4.5 Pemeriksaan Manipulasi Matematika**

Untuk memastikan bahwa kelompok matematika belajar matematika sepanjang tahun dan bahwa mahasiswa psikologi tidak, tes matematika dimasukkan. Ini terdiri dari 11

pertanyaan, tujuh di antaranya diambil dari subtes Perhitungan Woodcock-Johnson III, dua di antaranya adalah pertanyaan paling sulit pada tes diagnostik Universitas Stekom untuk sarjana matematika baru berdasarkan kinerja pada tahun 2008 dan 2009, dan dua terakhir yang didasarkan pada silabus gelar matematika tahun pertama. Tugas selengkapnya disajikan dalam Lampiran G.

## 5.5 PROSEDUR

Peserta mengambil bagian selama kuliah: mahasiswa matematika dalam satu kelompok dan mahasiswa psikologi di lain. RAPM selalu diselesaikan terlebih dahulu dengan batas waktu 15 menit. Sisa tugas diikuti dalam salah satu dari empat perintah penyeimbang persegi Latin yang ditugaskan secara acak kepada peserta:

1. Tes Matematika, Tugas Inferensi Kondisional Abstrak, Tugas Inferensi Kondisional Tematik, CRT
2. CRT, Tes Matematika, Tugas Inferensi Kondisional Abstrak, Tugas Inferensi Kondisional Tematik
3. Tugas Inferensi Kondisional Tematik, CRT, Tes Matematika, Tugas Inferensi Kondisional Abstrak
4. Tugas Inferensi Kondisional Abstrak, Tugas Inferensi Kondisional Tematik, CRT, Tes Matematika

## 5.6 HASIL

Dari 147 peserta yang mengikuti Time 1, 59 mahasiswa matematika dan 30 mahasiswa psikologi mengikuti kembali Time 2 dan dimasukkan dalam analisis. Tingginya angka putus sekolah mungkin disebabkan oleh sesi tes kedua yang berlangsung sesaat sebelum masa ujian, terutama pada kelompok psikologi. Selain itu, tidak semua peserta menyelesaikan semua tugas. Tabel 5.1 menunjukkan jumlah peserta dalam setiap kelompok yang menyelesaikan setiap tugas pada Waktu 2. Dalam semua analisis yang disajikan di bawah ini, data yang hilang dikeluarkan secara berpasangan untuk memaksimalkan kekuatan statistik.

**Tabel 5.1** Jumlah peserta di setiap kelompok yang menyelesaikan setiap tugas pada Waktu 2.

Tugas	Matematika	Psikolohi
RAPM	59	30
CRT	47	29
Inferensi Kondisional Abstrak	53	27
Inferensi Kondisional Tematik	49	28
Tes Matematika	51	30

Peserta yang kembali pada Waktu 2 memiliki skor yang lebih tinggi secara signifikan pada Waktu 1 RAPM,  $t(145) = 2.34, p = .021$ , Tes matematika Waktu 1,  $t(138) = 2.49, p = .014$ , Skor MCI Waktu 1 Abstrak,  $t(114) = 2.78, p = .006$ , dan skor MCI Tematik Time 1,  $t(116) = 2.19, p = .031$  dibandingkan mereka yang tidak kembali. Oleh karena itu sampel bias terhadap mahasiswa yang lebih mampu, tetapi akan selalu ada beberapa derajat bias ketika sampel dipilih sendiri. Tidak ada perbedaan antara mereka yang kembali dan mereka yang tidak dalam skor Time 1 CRT,  $t(145) = .17, p = .862$ . Yang penting, tidak ada interaksi yang signifikan antara Kelompok dan Pengembalian untuk RAPM, tes matematika atau skor MCI Tematik (semua

$p > .250$ ), menunjukkan bias tidak secara signifikan lebih besar dalam satu kelompok atau yang lain. Namun, terdapat interaksi yang signifikan antara skor MCI Abstrak Grup dan Return for Time 1,  $F(1,112) = 4,13, p = 0,044, \eta^2 = 0,04$ . Pada kelompok matematika, mereka yang kembali memiliki skor sedikit lebih tinggi daripada mereka yang tidak kembali,  $t(65) = 1,89, p = 0,063$ , sedangkan tidak ada perbedaan antara mereka yang kembali dan yang tidak kembali pada kelompok psikologi,  $t(47) = .99, p = .327$ . Skor rata-rata pada setiap tugas pada Waktu 1 untuk mereka yang kembali pada Waktu 2 dan mereka yang tidak kembali ditampilkan pada Tabel 5.2. Hasilnya dilaporkan dalam dua bagian: (i) analisis awal dan (ii) pengembangan skor interpretasi kondisional Abstrak dan Tematik.

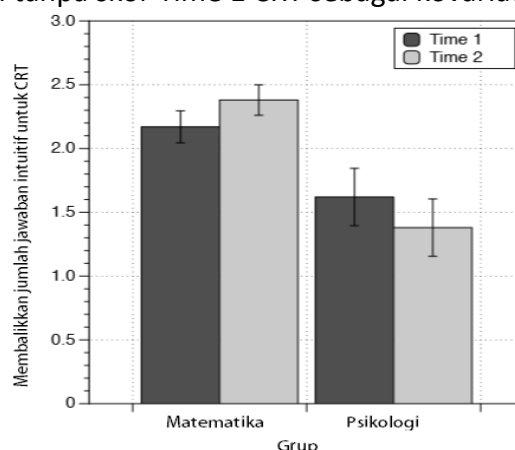
### 5.7 ANALISIS AWAL KOVARIAT

Kelompok matematika mendapat nilai RAPM yang signifikan lebih tinggi pada Time 1 ( $M = 11.1, SD = 2.97$ ) dibandingkan kelompok psikologi ( $M = 8.63, SD = 3.40$ ),  $t(87) = 3.56, p = .001, d = 0.77$ . Kelompok matematika juga mendapat skor yang lebih tinggi secara signifikan pada CRT pada Waktu 1 (jumlah jawaban intuitif terbalik,  $M = 2.12, SD = 0.86$ ) dibandingkan kelompok psikologi ( $M = 1.30, SD = 1.11$ ),  $U(72) = 327.50, z = 2,98, p = .003, r = .35$ .

Skor MCI abstrak pada Time 1 secara signifikan berkorelasi dengan skor Time 1 RAPM,  $r(63) = .33, p = .009$ , dan skor Time 1 CRT,  $r(63) = .26, p = .038$ . Akibatnya, skor Time 1 RAPM dan CRT digunakan sebagai kovariat dalam analisis selanjutnya dari indeks Interpretasi Kondisional Abstrak.

Skor MCI Tematik Time 1 secara signifikan berkorelasi dengan skor Time 1 RAPM,  $r(66) = .36, p = .003$ , dan Time 1 CRT,  $r(66) = .43, p < .001$ , mendukung penggunaan Waktu 1 skor RAPM dan CRT sebagai kovariat dalam analisis skor Interpretasi Kondisional Tematik.

Akhirnya, skor BBI pada Time 1 secara signifikan berkorelasi dengan skor Time 1 RAPM,  $r(66) = .26, p = .038$ , dan sedikit berkorelasi secara signifikan dengan skor Time 1 CRT,  $r(66) = .24, p = 0,058$ . Analisis BBI dilaporkan di bawah ini dengan skor Time 1 RAPM sebagai kovariat dan keduanya dengan dan tanpa skor Time 1 CRT sebagai kovariat.



**Gambar 5.3:** Interaksi antara Grup dan Waktu pada jumlah (terbalik) jawaban intuitif yang diberikan kepada CRT (di mana skor yang lebih tinggi menunjukkan kinerja yang lebih baik, bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata).

Meskipun kedua kelompok sedikit meningkatkan skor RAPM mereka sepanjang tahun, efek interaksi Grup  $\times$  Waktu tidak mendekati signifikan,  $F(1,80) < 1$ . Sekali lagi, kedua kelompok sedikit meningkatkan skor CRT mereka sepanjang tahun tetapi Grup  $\times$  Pengaruh

interaksi waktu tidak mendekati signifikansi,  $F(1,62) = 1,95, p = 0,167, \eta^2 = 0,031$  (ditunjukkan pada Gambar 5.3).

### 5.7.1 Pemeriksaan Manipulasi

Perubahan nilai tes matematika dianalisis dengan ANOVA 2x2 dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan satu faktor antar mata pelajaran: Kelompok (matematika, psikologi). Ada interaksi yang signifikan,  $F(1,76) = 8,24, p = 0,005$ , yang menunjukkan bahwa kelompok matematika meningkat ke tingkat yang lebih besar daripada kelompok psikologi, ditunjukkan pada Gambar 5.4. Peningkatan kelompok matematika dari waktu ke waktu dikonfirmasi oleh perbandingan yang direncanakan dari skor Waktu 1 dan 2,  $t(48) = 3,70, p = .001, d = .70$ . Ini menunjukkan bahwa sebagai kelompok mereka terlibat dan belajar dari tahun mereka belajar matematika dan manipulasi kuasi berhasil.

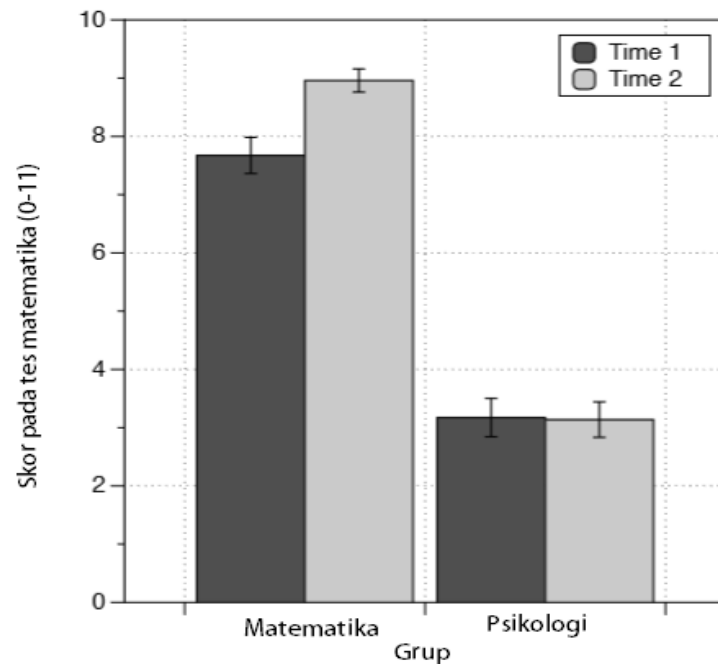
**Tabel 5.2:** Nilai rata-rata pada setiap tugas pada Waktu 1 untuk mereka yang melakukan dan tidak kembali untuk mengambil bagian pada Waktu 2 dengan standar deviasi dalam tanda kurung.

Tugas Maksimal	Kembali ke waktu 2	Absen Waktu 2
RAPM (18)	10.27 (3.29)	8.91 (3.64)
CRT (3)	2.06 (1.02)	2.07 (1.03)
Abstrak MCI (32)	20.58 (4.51)	18.43 (3.43)
Tematik MCI (16)	11.16 (2.91)	9.93 (3.02)
Tes Matematika (11)	6.23 (2.94)	4.98 (2.82)

## 5.8 PENGEMBANGAN KEAHLIAN PENALARAN

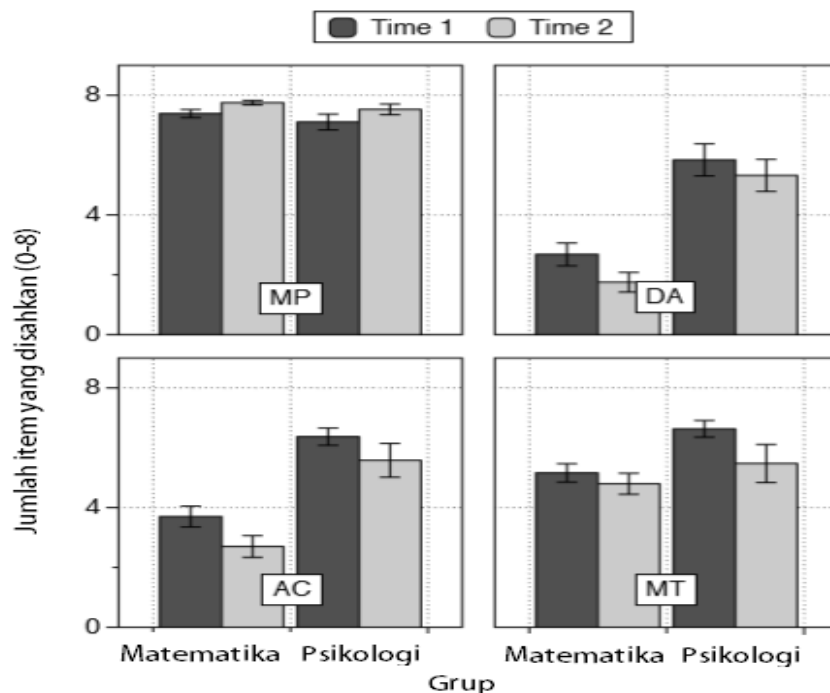
### 5.8.1 Inferensi Kondisional Abstrak.

Tingkat pengesahan. Tingkat pengesahan dari setiap jenis inferensi dianalisis dengan ANOVA 2x4x2 dengan dua faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan Jenis Inferensi (MP, DA, AC, MT), satu faktor antar mata pelajaran : Kelompok (matematika dan psikologi), dan dua kovariat: Waktu 1 RAPM dan Waktu 1 CRT. Ini menunjukkan interaksi tiga arah yang sedikit signifikan,  $F(3,177) = 2.58, p = .055, \eta^2 = .04$ , (lihat Gambar 6.5). Rata-rata dan simpangan baku untuk interaksi ini ditampilkan pada Tabel 6.3. Pada Waktu 2 mahasiswa matematika mendukung lebih banyak inferensi MP,  $t(43) = 2.33, p = .025, d = 0.50$ , dan lebih sedikit DA,  $t(43) = 3.37, p = .002, d = 0.40$  dan inferensi AC,  $t(43) = 3.30, p = .002, d = 0.42$  dibandingkan dengan Waktu 1. Mahasiswa matematika tidak mengalami perubahan dalam tingkat dukungan inferensi MT,  $t(43) = 1.19, p = .243, d = 0.16$ . Sebaliknya, kelompok psikologi tidak menunjukkan perbedaan yang signifikan antara skor Time 1 dan Time 2 untuk inferensi apapun (semua  $ps > .160$ ), meskipun ada sedikit penurunan signifikan dalam jumlah inferensi MT yang didukung,  $t(18) = 1,97, p = .065, d = 0.54$ .



**Gambar 5.4:** Interaksi antara Kelompok dan Waktu pada nilai tes matematika (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar dari mean).

Peningkatan dukungan MP bersama dengan penurunan dukungan DA dan AC konsisten dengan interpretasi yang lebih material atau cacat dari kondisi tersebut. Untuk menyelidiki ini secara formal, setiap indeks interpretasi dianalisis dengan ANOVA 2x2 dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan satu faktor antara mata pelajaran: Kelompok (matematika, psikologi). Nilai indeks rata-rata untuk setiap kelompok pada setiap titik waktu ditunjukkan pada Gambar 5.6.



**Gambar 5.5:** Rata-rata tingkat dukungan untuk masing-masing dari empat kesimpulan di setiap kelompok pada Waktu 1 dan Waktu 2 pada Tugas Inferensi Kondisional Abstrak (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata).

**Tabel 5.3:** Jumlah rata-rata item yang didukung pada Tugas Inferensi Kondisional Abstrak menurut jenis Inferensi, Kelompok dan titik Waktu dengan standar deviasi dalam tanda kurung.

Intervensi	Grup	Waktu 1	Waktu 2
MP	Matematika	7.39 (0.89)	7.75
	B. Inggris	7.11 (1.15)	7.53 (0.77)
DA	Matematika	2.68 (2.52)	1.75 (2.18)
	B. Inggris	5.84 (2.34)	5.32 (2.33)
AC	Matematika	3.70 (2.31)	2.70 (2.41)
	B. Inggris	6.37 (1.26)	5.58 (2.46)
MT	Matematika	5.16 (2.05)	4.80 (2.32)
	B. Inggris	6.63 (1.21)	5.47 (2.78)

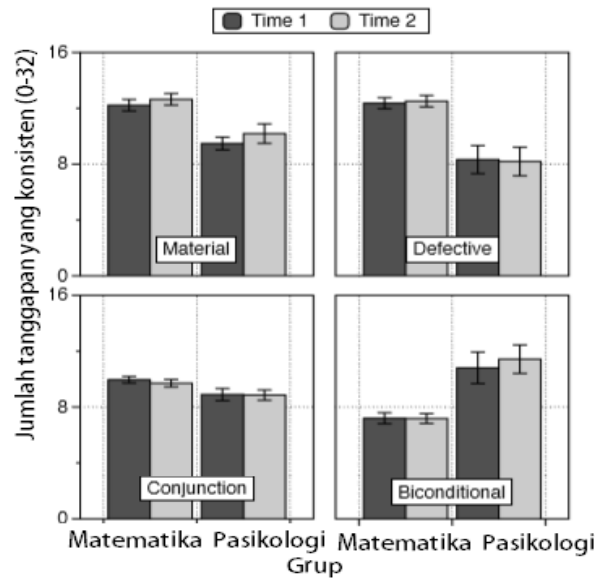
### 5.8.2 Interpretasi

Untuk mengevaluasi perubahan kecenderungan masing-masing kelompok terhadap empat interpretasi kondisi abstrak, dilakukan empat ANOVA 2x2, satu untuk setiap interpretasi, masing-masing dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun), satu antara -faktor mata pelajaran: Kelompok (matematika, psikologi) dan dua kovariat: RAPM dan CRT. Rerata indeks interpretasi masing-masing kelompok pada setiap titik waktu disajikan pada Tabel 5.4 dan Gambar 5.6.

Untuk MCI, tidak ada interaksi yang signifikan,  $F(1,59) = 2.20, p = .143, \eta^2 = .04$ , tetapi ada kecenderungan arah yang diprediksi oleh TFD. Uji-t sampel berpasangan menunjukkan bahwa MCI kelompok matematika secara signifikan lebih tinggi pada Waktu 2 ( $M=24,09, SD=4,10$ ) daripada pada Waktu 1 ( $M=22,16, SD=4,24$ ),  $t(43) = 3,68, p = .001, d = 0,55$ , sedangkan MCI kelompok psikologi tidak berubah dari waktu ke waktu,  $t(18) = .70, p = .495, d = 0,18$  (Waktu 1:  $M=17,53, SD=3,61$ , Waktu 2:  $M= 18.11, SD=2.69$ ). Selanjutnya, MCI kelompok matematika secara signifikan lebih tinggi daripada kelompok psikologi pada Time 1,  $t(67) = 4.23, p < .001, d = 1.17$ , dan pada Time 2,  $t(78) = 6.71, p < .001, d = 1,72$ .

**Tabel 5.4:** Nilai indeks rata-rata untuk setiap interpretasi pernyataan kondisional abstrak pada Waktu 1 dan Waktu 2 pada setiap kelompok (standar deviasi dalam tanda kurung).

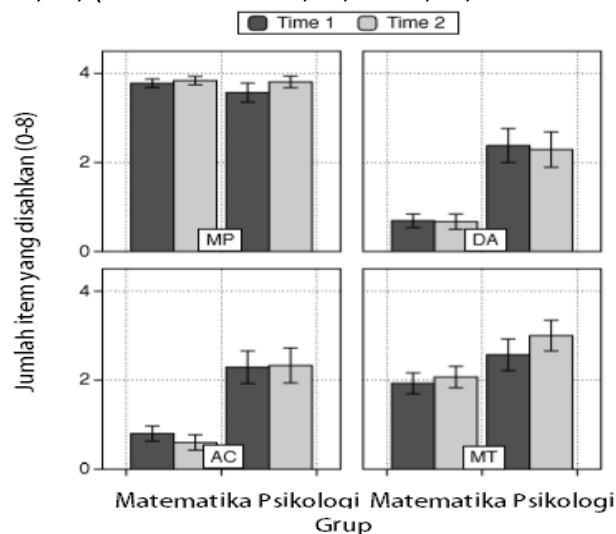
Grup	Waktu	Materi	Defektif	Bikondisional	Konjungtif
Matematika	Waktu 1	22.16 (4.24)	19.84 (5.58)	18.93 (5.67)	19.25 (2.92)
	Waktu 2	24.09 (4.10)	22.50 (5.98)	17.00 (5.61)	19.91 (2.76)
Kontrol	Waktu 1	17.53 (3.61)	12.26 (4.07)	25.95 (3.88)	17.00 (2.87)
	Waktu 2	18.11 (2.69)	15.16 (7.34)	23.89 (7.20)	18.32 (3.16)



**Gambar 5.6:** Interaksi antara Grup dan Waktu pada indeks Inferensi Kondisional Abstrak (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata).

Pola hasil serupa muncul untuk DCI. Meskipun tidak ada interaksi yang signifikan,  $F < 1$ , uji-t sampel berpasangan menunjukkan bahwa DCI kelompok matematika secara signifikan lebih tinggi pada Waktu 2 ( $M=22,50$ ,  $SD=5,98$ ) daripada pada Waktu 1 ( $M=19,84$ ,  $SD=5,58$ ),  $t(43) = 4.35$ ,  $p < .001$ ,  $d = 0.66$ , sedangkan DCI kelompok psikologi hanya sedikit lebih tinggi pada Time 2 ( $M=15.16$ ,  $SD=7.34$ ), dibandingkan pada Time 1 ( $M=12.26$ ,  $SD= 4,07$ ),  $t(18) = 2,05$ ,  $p = 0,056$ ,  $d = 0,49$ . DCI kelompok matematika secara signifikan lebih tinggi daripada kelompok psikologi pada Time 1,  $t(67) = 5.28$ ,  $p < .001$ ,  $d = 1.55$ , dan Time 2,  $t(78) = 4.98$ ,  $p < .001$ ,  $d = 1.09$ .

Untuk CCI, tidak ada interaksi yang signifikan antara Waktu dan Grup,  $F(1,59) = 1,50$ ,  $p = .225$ ,  $\eta^2 = .025$ . Nilai kelompok matematika tidak berubah dari waktu ke waktu,  $t(43) = 1.45$ ,  $p = .155$ ,  $d = .23$ , sedangkan IPK kelompok psikologi meningkat sedikit dari waktu ke waktu,  $t(18) = 1.99$ ,  $p = .062$ ,  $d = 0,44$ , (Waktu 1:  $M=17,00$ ,  $SD=2,87$ , Waktu 2:  $M=18,32$ ,  $SD=3,16$ ).



**Gambar 5.7:** Rata-rata tingkat dukungan untuk masing-masing dari empat kesimpulan di setiap kelompok pada Waktu 1 dan Waktu 2 pada Tugas Inferensi Kondisional Tematik (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata).



Analisis skor BCI juga menunjukkan tidak ada interaksi yang signifikan,  $F(1,59) < 1$ , tetapi uji-t sampel berpasangan menunjukkan bahwa BCI kelompok matematika menurun dari waktu ke waktu,  $t(43) = 3,07, p = 0,004, d = 0.46$ , (Waktu 1:  $M=18.93, SD=5.67$ , Waktu 2:  $M=17.00, SD=5.61$ ), sedangkan BCI kelompok psikologi tidak berubah,  $t(18) = 1.46, p = .161, d = 0.36$ .

Kekuatan untuk analisis penelitian rendah, karena angka putus sekolah yang tinggi secara tak terduga dalam kelompok psikologi menghasilkan 19 peserta psikologi yang menyelesaikan tugas Inferensi kondisional abstrak dan melakukan kovariat pada kedua titik waktu<sup>4</sup>. Uji-t yang disajikan di atas menunjukkan bahwa ada tren ke arah yang diprediksi oleh hasil level AS dalam kelompok matematika. Namun, cara yang disajikan pada Tabel 5.4 menunjukkan bahwa kelompok psikologi menunjukkan pola perubahan yang serupa, meskipun tidak secara signifikan dan dengan ukuran efek yang lebih kecil.

### 5.8.3 Inferensi Kondisional Tematik.

**Tingkat pengesahan** dari setiap jenis inferensi dianalisis dengan ANOVA 2x4x2 dengan dua faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan Jenis Inferensi (MP, DA, AC, MT), satu faktor antar mata pelajaran : Kelompok (matematika dan psikologi), dan dua kovariat: Waktu 1 RAPM dan Waktu 1 CRT. Tidak ada interaksi tiga arah yang signifikan,  $F < 1$ , (lihat Gambar 5.7). Rata-rata dan simpangan baku untuk interaksi ini ditampilkan pada Tabel 5.5. Tidak ada kelompok yang berubah dalam tingkat dukungan mereka dari salah satu jenis inferensi dari waktu ke waktu (semua  $ps > .25$ ).

### 5.8.4 Interpretasi

Selanjutnya, kecenderungan masing-masing kelompok terhadap empat interpretasi kondisional tematik diselidiki. Empat ANOVA 2x2 dilakukan, masing-masing dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun), satu faktor antara mata pelajaran: Kelompok (matematika, psikologi) dan dua kovariat: RAPM dan CRT. Rerata indeks interpretasi masing-masing kelompok pada setiap titik waktu disajikan pada Tabel 5.6 dan Gambar 5.8. Tidak ada interaksi yang signifikan antara Waktu dan Grup untuk salah satu indeks interpretasi, semua  $ps > .330$ . Lebih lanjut, tidak seperti pada tugas Inferensi Kondisional Abstrak, empat nilai indeks mahasiswa matematika tidak berubah dari waktu ke waktu pada versi tematik, semua  $ps > .235$ , dan begitu pula nilai mahasiswa psikologi, semua  $ps > .175$ .

**Tabel 5.5:** Jumlah rata-rata item yang didukung pada Tugas Inferensi Kondisional Tematik menurut jenis Inferensi, Kelompok dan Titik waktu dengan standar deviasi dalam tanda kurung.

Intervensi	Grup	Waktu 1	Waktu 2
MP	Matematika	3.78 (0.64)	3.84 (0.64)
	Psikologi	3.57 (0.98)	3.81 (0.60)
DA	Matematika	0.69 (1.02)	0.67 (1.15)
	Psikologi	2.38 (1.75)	2.29 (1.82)
AC	Matematika	0.80 (1.14)	0.60 (1.16)
	Psikologi	2.29 (1.68)	2.33 (1.80)
MT	Matematika	1.98 (1.57)	2.07 (1.63)
	Psikologi	2.57 (1.63)	3.00 (1.58)

<sup>4</sup> Kelompok kedua mahasiswa saat ini mengambil bagian dalam studi dan analisis yang lebih kuat akan dilakukan setelah Mei 2013.

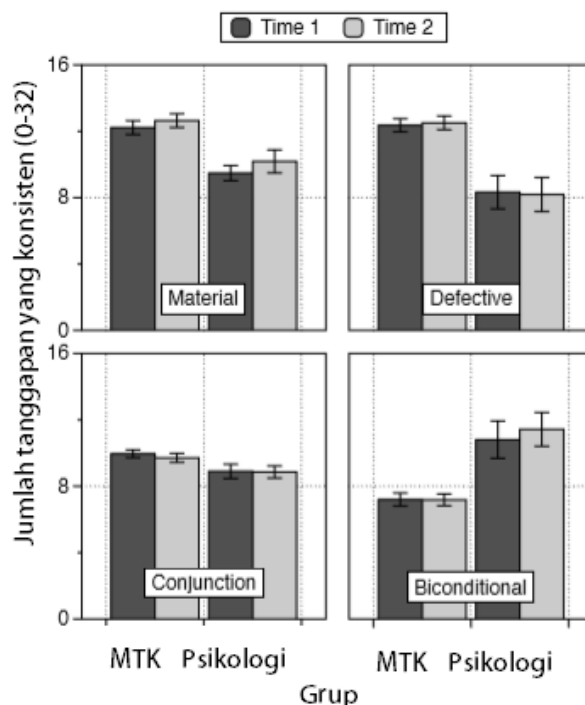
Seperti halnya tugas Inferensi, Abstrak Kondisional, kekuatan untuk menganalisis tugas Inferensi Kondisional Tematik rendah karena tingkat putus sekolah yang tinggi dalam kelompok psikologi. Namun, dalam kasus ini bahkan uji-t berpasangan yang membandingkan indeks interpretasi kelompok matematika di dua titik waktu yang signifikan.

Akhirnya, skor BBI untuk setiap kelompok pada setiap titik waktu dianalisis dengan ANOVA  $2 \times 2$  dengan satu faktor dalam mata pelajaran: Waktu (awal dan akhir tahun) dan satu faktor antar mata pelajaran: Kelompok (matematika, psikologi), dan dua kovariat, RAPM dan CRT. Ini menunjukkan tidak ada efek utama Grup,  $F(1,62) < 1$ , tidak ada efek utama Waktu  $F(1,62) < 1$ , dan tidak ada interaksi,  $F(1,62) < 1$  (yang tetap tidak signifikan tanpa CRT sebagai kovariat,  $p = 0,482$ , lihat Tabel 6.7 untuk mean dan standar deviasi). Paired t-tests menunjukkan bahwa tidak ada kelompok BBI yang berubah secara signifikan dari waktu ke waktu (keduanya  $ps > .290$ ).

Karena tidak ada interaksi yang signifikan antara Grup dan Waktu untuk salah satu ukuran hasil, tidak ada kebutuhan untuk analisis regresi yang memprediksi perubahan dari RAPM, CRT dan Grup.

**Tabel 5.6:** Nilai indeks rata-rata untuk setiap interpretasi pernyataan kondisional tematik pada Waktu 1 dan Waktu 2 pada masing-masing kelompok (standar deviasi dalam tanda kurung).

Grup	Waktu	Materi	Defektif	Bikondisional	Konjungtif
Matematika	Waktu 1	12.22 (2.79)	12.36 (2.66)	7.20 (2.64)	9.56 (1.59)
	Waktu 2	12.64 (2.78)	12.51 (2.76)	7.18 (2.38)	9.71 (1.80)
Kontrol	Waktu 1	9.48 (2.06)	8.33 (4.61)	10.81 (5.16)	8.90 (1.97)
	Waktu 2	10.19 (3.19)	8.19 (4.69)	11.43 (4.65)	8.86 (1.68)



**Gambar 5.8:** Interaksi antara Grup dan Waktu pada indeks interpretasi Inferensi Kondisional Tematik (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata).

**Tabel 5.7:** Rata-rata skor indeks bias keyakinan untuk setiap kelompok pada Waktu 1 dan Waktu 2 (standar deviasi dalam tanda kurung).

	Matematika	Psikologi
Waktu 1	.31 (1.47)	.62 (1.24)
Waktu 2	.02 (1.20)	.38 (1.02)

## 5.9 DISKUSI

Bab 5 ini bertujuan untuk mengetahui perkembangan kemampuan penalaran pada mahasiswa sarjana matematika. Mahasiswa matematika dibandingkan dengan mahasiswa psikologi dalam penyelidikan longitudinal perilaku penalaran kondisional abstrak dan tematik. Hipopenelitian 1 meramalkan bahwa mahasiswa matematika akan menjadi lebih cacat, dan pada tingkat lebih rendah lebih banyak materi, dalam interpretasi mereka terhadap pernyataan kondisional abstrak. Hipopenelitian ini tidak didukung oleh serangkaian analisis ANOVA, tetapi uji-t berpasangan menunjukkan bahwa penalaran kondisional abstrak mahasiswa matematika menjadi lebih cacat ( $d = 0,66$ , yang mirip dengan perubahan pada mahasiswa matematika tingkat AS,  $d = 0,88$ ), lebih banyak materi ( $d = 0,55$ , dibandingkan dengan  $d = 0,49$  pada mahasiswa matematika tingkat AS), dan lebih sedikit bikondisional dari waktu ke waktu ( $d = 0,46$ , dibandingkan dengan  $d = 0,51$  pada mahasiswa matematika tingkat AS), sedangkan psikologi perilaku mahasiswa tidak berubah. Hal ini konsisten dengan Hipopenelitian 1 dan dengan hasil dari mahasiswa tingkat AS di Bab 5, tetapi kurangnya signifikansi statistik dalam analisis interaksi mencegah kesimpulan yang pasti dari ditarik. Analisis menderita dari kekuatan statistik yang rendah karena ukuran sampel yang kecil dalam kelompok psikologi, dan ini mungkin telah mencegah analisis ANOVA mencapai signifikansi.

Tidak ada prediksi yang jelas untuk tugas *Inferensi Kondisional Tematik*. Bab 5 menunjukkan tidak ada perubahan dalam tugas Silogisme Tematik pada mahasiswa matematika AS, dan tidak jelas apakah perbedaan dengan perubahan inferensi kondisional abstrak dalam penelitian itu karena aspek tematik tugas atau aspek silogistik. Dalam penelitian ini, penalaran mahasiswa matematika dengan kondisional abstrak menunjukkan tren ke arah yang diprediksi, tetapi tidak menunjukkan tren dalam tugas Inferensi Kondisional Tematik, meskipun item memiliki struktur yang sama. Ini menunjukkan bahwa konteks mungkin menjadi masalahnya. Mungkin kasus bahwa belajar matematika hanya mempengaruhi kemampuan seseorang untuk bernalar tentang pernyataan kondisional abstrak, yang masuk akal mengingat banyak dari matematika itu sendiri bersifat abstrak. Jika demikian halnya, ada implikasi penting bagi Teori Disiplin Formal. Pernyataan seperti yang dibuat oleh Oakley, bahwa “studi matematika tidak dapat digantikan oleh aktivitas lain apa pun yang akan melatih dan mengembangkan kemampuan logika murni manusia ke tingkat rasionalitas yang sama” (1949, p.19), dan Plato, bahwa “kita harus berusaha untuk membujuk mereka yang akan menjadi orang-orang utama negara kita untuk pergi dan belajar aritmatika” (375B.C/2003, p. 256), tampaknya menyiratkan pengaruh yang lebih umum dari belajar matematika pada keterampilan penalaran, pengaruh yang akan meluas ke penalaran rasional sehari-hari. Jika matematika sebenarnya hanya mempengaruhi perilaku penalaran kondisional abstrak, maka dampaknya pada penalaran sehari-hari tentu sangat terbatas.

Isu lain yang diangkat pada bab 4 adalah kerancuan antara pelajaran matematika dan IPA di tingkat A. Itu tidak mungkin untuk menyelidiki pengaruh belajar matematika secara independen dari mata pelajaran IPA, tetapi tidak ada korelasi antara jumlah mata pelajaran

IPA yang dipelajari dan perubahan DCI dalam kelompok matematika. Partisipan sarjana dalam penelitian ini semuanya mengambil gelar kehormatan tunggal, artinya mahasiswa matematika hanya mempelajari matematika, dan mereka masih menunjukkan kecenderungan menjadi lebih material dan cacat dalam penalaran mereka. Hal ini menunjukkan bahwa matematika saja dikaitkan dengan perubahan dalam penalaran kondisional, meskipun efeknya mungkin masih dibesar-besarkan dengan mempelajari mata pelajaran ilmu lainnya secara bersamaan.

Masalah terakhir yang ingin diselesaikan oleh penelitian ini adalah perbedaan temuan antara Lehman dan Nisbett (1990) dan Inglis dan Simpson (2009a). Lehman dan Nisbett (1990) menemukan bahwa jumlah modul matematika yang diambil oleh mahasiswa sarjana di universitas AS berkorelasi dengan tingkat peningkatan mereka pada tugas penalaran kondisional. Inglis dan Simpson (2009a), bagaimanapun, gagal untuk menemukan perbaikan apapun pada tugas inferensi kondisional abstrak di sarjana matematika di universitas Indonesia. Data yang disajikan di sini menambahkan (sejumlah kecil) bobot temuan Lehman dan Nisbett (1990), menunjukkan bahwa ada hubungan antara matematika tingkat sarjana dan pengembangan keterampilan penalaran kondisional. Alasan tidak adanya perubahan atau tren yang ditemukan dalam studi Inglis dan Simpson (2009a) tidak jelas – ukuran ketergantungan mereka dan tingkat studi peserta sama seperti dalam penelitian ini. Mungkin modul yang dipelajari oleh peserta kami dan mereka berbeda dan terkait secara berbeda dengan keterampilan penalaran. Kemungkinan lain adalah ada perubahan dalam interpretasi yang salah, yang tidak diselidiki oleh Inglis dan Simpson (2009a). Penelitian lebih lanjut yang membandingkan pengaruh mempelajari kurikulum matematika yang berbeda, baik di tingkat sarjana dan pendidikan lainnya, pada pengembangan keterampilan penalaran kondisional akan berguna untuk memperjelas masalah ini.

Mengenai mekanisme potensial peningkatan mahasiswa matematika dalam keterampilan penalaran kondisional abstrak, data yang disajikan dalam bab ini sayangnya tidak banyak membantu. Tren yang dilaporkan konsisten dengan hipotesis yang diuraikan dalam Bab 5, bahwa belajar matematika mengajarkan mahasiswa untuk mengasumsikan  $p$  dan alasan tentang  $q$ , sehingga membuat penalaran mereka dari pernyataan kondisional lebih cacat. Namun, kurangnya interaksi yang signifikan antara Waktu dan Grup untuk DCI atau indeks lainnya berarti bahwa hasilnya hanya sugestif, dan tidak mungkin untuk menguji pengaruh perubahan RAPM dan CRT pada perubahan perilaku penalaran.

## 5.10 RINGKASAN DAN TEMUAN BARU

Studi yang disajikan di sini telah membuat tiga kontribusi tentatif untuk pemahaman kita tentang Teori Disiplin Formal:

1. Belajar matematika di tingkat sarjana dikaitkan dengan kecenderungan penalaran yang lebih cacat, lebih banyak materi, dan kurang bikondisional.
2. Belajar matematika secara terpisah dari mata pelajaran ilmu dikaitkan dengan perubahan serupa dengan yang ditemukan pada mahasiswa tingkat AS yang belajar matematika bersama dengan mata pelajaran ilmu.
3. Perubahan hanya berlaku untuk masalah penalaran kondisional abstrak, bukan masalah tematik.

## BAB 6

### FAKTOR BAHASA MATEMATIKA

#### 'JIKA KEMUDIAN' DAN 'HANYA JIKA'

##### 6.1 PENDAHULUAN

Tujuan dari bab ini adalah untuk menetapkan apakah mahasiswa matematika memiliki pemahaman yang mendalam tentang logika kondisional, yaitu apakah mereka bernalar berdasarkan struktur logis dari kondisional, atau apakah mereka dipengaruhi oleh fitur permukaan? Bab 4 dan 5 mengemukakan bahwa perkembangan penalaran mahasiswa matematika terbatas pada inferensi kondisional abstrak, tidak meluas ke inferensi kondisional tematik atau silogisme tematik. Pertanyaan yang diangkat di sini adalah sejauh mana penalaran mahasiswa matematika digeneralisasikan dalam inferensi kondisional abstrak. Tugas inferensi kondisional yang digunakan dalam studi lain dari penelitian ini menyajikan pernyataan kondisional kepada peserta dalam bentuk 'jika  $p$  maka  $q$ '. Cara yang setara secara logis untuk mengungkapkan pernyataan ini adalah ' $p$  hanya jika  $q$ ' (Evans, 1977), seperti yang ditunjukkan dalam bentuk tabel kebenaran pada Tabel 6.1. Jika mahasiswa matematika memiliki pemahaman yang mendalam tentang logika kondisional, maka mereka harus memperlakukan pernyataan 'jika maka' (IT) dan 'hanya jika' (OI) dengan cara yang sama. Ini tidak terjadi dengan mahasiswa non-matematika dalam studi oleh Evans (1977).

Evans (1977) disajikan mahasiswa sarjana, yang tidak khusus dari kursus matematika, dengan versi 16 item tugas inferensi kondisional. Setengah dari pesertanya melihat item yang diutarakan 'jika  $p$  maka  $q$ ' dan separuh lainnya melihatnya sebagai ' $p$  hanya jika  $q$ '. Evans berhipotesis perbedaan interpretasi antara IT dan aturan OI, berdasarkan prinsip-prinsip kebutuhan dan kecukupan. Dalam implikasi material, anteseden cukup untuk konsekuen dan konsekuen diperlukan untuk anteseden. Namun, kedua prinsip ini ditekankan secara berbeda dalam aturan IT dan OI. Menggunakan contoh Evans, aturan IT 'Jika dia seorang polisi maka tingginya lebih dari 5 kaki 9 inci' tampaknya menekankan kecukupan dari pendahulunya. Di sisi lain, aturan OI yang setara 'Dia adalah seorang polisi hanya jika tingginya lebih dari 5 kaki 9 inci', tampaknya menekankan perlunya konsekuensi (Evans, 1977, hlm. 300). Berdasarkan perbedaan ini, Evans menghipotesiskan dua hal: pertama, bahwa lebih banyak kesimpulan MP akan dibuat pada aturan TI daripada aturan OI, karena premis minor dalam MP menguatkan anteseden, yang efisiensinya ditekankan dalam aturan TI, dan kedua, bahwa lebih banyak MT kesimpulan akan dibuat pada aturan OI daripada aturan IT, karena premis minor di MT meniadakan konsekuensi, yang kebutuhannya ditekankan dalam aturan OI (ketika konsekuensi yang diperlukan dinegasikan, lebih jelas dari aturan OI bahwa anteseden juga harus dinegasikan) .

**Tabel 6.1:** Tabel kebenaran untuk pernyataan kondisional ' $p$  hanya jika  $q$ ' dan ' $p$  maka  $q$ '.

$p$	$q$	$p$ hanya jika $q$	jika $p$ maka $q$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	T

Hipopenelitian ini didukung oleh data. Inferensi MP didukung 100% sepanjang waktu dalam kondisi TI, dibandingkan dengan 76% dalam kondisi OI. Inferensi MT, di sisi lain, didukung 42% waktu dalam kondisi TI dibandingkan dengan 59% waktu dalam kondisi OI. Perbedaan tak terduga yang ditemukan oleh Evans (1977) adalah pada inferensi AC. Evans tidak memprediksi perbedaan antara interpretasi TI dan OI dari inferensi AC, tetapi menemukan tingkat dukungan yang lebih tinggi dalam kondisi OI, sebesar 84%, dibandingkan dengan kondisi TI, sebesar 67%. Evans menduga bahwa ini bisa jadi karena peserta melakukan konversi, di mana mereka mengambil 'p hanya jika q' berarti 'jika q maka p'; pembacaan IT dengan anteseden dan konsekuen dibalik. Dalam hal ini, inferensi OI AC akan setara dengan inferensi MP IT, yang dapat menjelaskan tingkat dukungan yang tinggi dari inferensi AC yang tidak valid dalam kondisi OI. Seorang MP OI akan menjadi IT AC, OI DA menjadi IT MT, dan OI MT akan menjadi IT DA. Ketika direklasifikasi dengan cara ini, tingkat dukungan OI menjadi lebih mirip dengan tingkat dukungan TI pada semua kecuali OI MT/IT DA (lihat Tabel 6.2), sehingga hipopenelitian konversi dapat menjelaskan perbedaan interpretasi.

**Tabel 6.2:** Persentase tingkat dukungan untuk setiap inferensi TI dan inferensi OI yang dikonversi setara dalam studi Evans (1977).

Intervensi IT	Dukungan IT	Dukungan OI	Intervensi OI
MP	100	84	AC
DA	38	59	MT
AC	67	76	MP
MT	42	38	DA

Dalam studi yang disajikan di bawah ini, tujuannya adalah untuk melihat apakah mahasiswa matematika tingkat lanjut akan menunjukkan perbedaan interpretasi yang sama yang ditemukan oleh Evans (1977), atau apakah mereka akan merespons logika kondisional dan menafsirkan pernyataan TI dan OI dengan cara yang sama. , mengingat bahwa mereka secara logis setara. TFD akan memprediksi skenario kedua; ini menunjukkan bahwa mempelajari matematika "mengembangkan kemampuan logika murni manusia" (Oakley, 1949, hlm. 19), yang diambil di sini berarti bahwa mereka memperoleh pemahaman yang mendalam tentang logika yang bebas dari pengaruh interferensi tingkat permukaan.

Namun, jika mahasiswa matematika menjadi terbiasa dengan inferensi maju, seperti 'jika p maka q', dan belajar mengasumsikan p, mereka akan merespon secara berbeda terhadap inferensi 'mundur' yang tampaknya 'p hanya jika q'. Dalam hal ini, ajakan itu bukan untuk menganggap p melainkan untuk mempertanyakannya. Tidak jelas apa interpretasi yang lebih disukai dari pernyataan OI dalam kasus ini, tetapi interpretasi yang rusak mungkin akan secara signifikan lebih rendah dalam kondisi OI daripada dalam kondisi TI. Mungkin diharapkan bahwa inferensi AC akan didukung secara signifikan lebih banyak dalam kondisi OI daripada dalam kondisi TI; kondisi OI menekankan ketidakpastian q, tetapi kebutuhan p ketika q benar. Oleh karena itu inferensi AC, di mana q dikonfirmasi, mungkin akan menyebabkan tingkat dukungan yang tinggi. Ini adalah pola yang ditemukan Evans (1977) dengan mahasiswa non-matematika (dijelaskan sebagai konversi dari OI AC ke IT MP). Jika mahasiswa matematika melakukan penalaran secara salah karena mereka telah terbiasa dengan inferensi IT 'maju', maka dalam kasus inferensi 'mundur' (OI) kita akan mengharapkan mereka untuk bernalar

dengan cara yang sama seperti mahasiswa non-matematika: dengan dukungan yang lebih tinggi dari AC.

## 6.2 METODE

### 6.2.1 Partisipan

Partisipan adalah 61 mahasiswa sarjana matematika tahun ketiga di Stekom University. Mereka mengambil bagian secara sukarela tanpa dibayar selama kuliah tentang Statistik Terapan dan kemudian menganalisis data yang dianonimkan sebagai bagian dari sesi lab SPSS.

### 6.2.2 Desain

Studi ini mengikuti desain eksperimental antara subjek dengan dua kondisi: 'jika kemudian' dan 'hanya jika' ungkapan tugas inferensi kondisional. Tiga puluh peserta ditugaskan untuk kondisi 'jika kemudian' dan 31 untuk kondisi 'hanya jika'.

### 6.2.3 Pengukuran

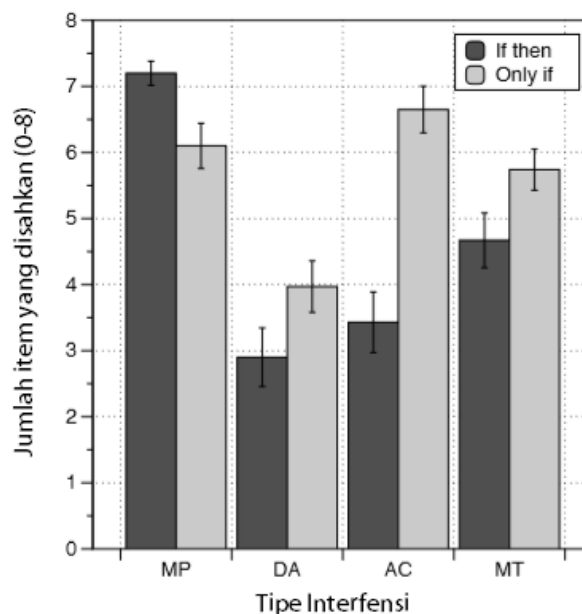
Peserta menyelesaikan tugas inferensi kondisional 32 item dengan pernyataan kondisional yang diutarakan sebagai 'jika p maka q' (kondisi IT) atau sebagai 'p hanya jika q' (kondisi OI). Huruf-huruf dan angka-angka yang digunakan dalam soal itu identik di seluruh kondisi, seperti halnya instruksi tugas. Urutan item diacak antara peserta.

### 6.2.4 Prosedur

Peserta mengambil bagian selama kuliah. Mereka diberitahu bahwa mereka akan diberikan tugas penalaran dan bahwa mereka akan menganalisis data dalam sesi lab SPSS berikutnya. Peserta diminta untuk bekerja sendiri dan dalam diam dan tidak diberitahu bahwa ada dua versi tugas. Buklet tugas dibagikan kepada peserta dengan syarat secara bergantian sehingga syarat yang diterima peserta hanya bergantung pada urutan duduk di ruang kuliah.

### 6.2.5 Hasil

Hasilnya disajikan dalam dua bagian. Pertama, analisis tingkat dukungan untuk setiap kesimpulan disajikan untuk perbandingan dengan penelitian Evans (1977). Kedua, disajikan analisis indeks interpretasi pada setiap kondisi.

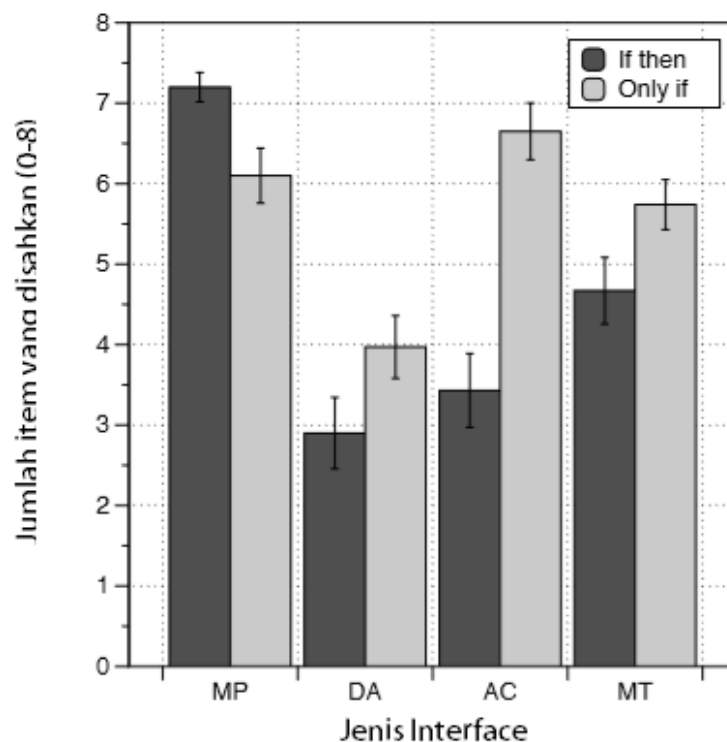


**Tabel 6.3:** Jumlah rata-rata item yang didukung (dari 8) untuk masing-masing dari empat kesimpulan berdasarkan kondisi. Standar deviasi dalam tanda kurung.

#### 6.4 TARIF PENGESAHAN

Rata-rata tingkat dukungan menjadi sasaran 2 (kondisi: IT, OI)  $\times$  4 (tipe inferensi: MP, DA, AC, MT) analisis varians (ANOVA). Ini mengungkapkan efek utama yang signifikan dari tipe inferensi,  $F(3,177) = 35.81, p < .001, \eta^2 = .38$ , dengan item MP paling sering didukung ( $M = 6.64, SD = 1.61$ ), diikuti oleh item MT ( $M = 5.21, SD = 2.07$ ), diikuti oleh item AC ( $M = 5.07, SD = 2.76$ ), dengan item DA paling jarang didukung ( $M = 3.44, SD = 2.34$ ). Ada juga efek utama yang signifikan dari kondisi,  $F(1,59) = 8.74, p = .004, \eta^2 = .13$ , dengan lebih banyak item yang didukung dalam kondisi OI ( $M = 22.45, SD = 5.27$ ) daripada IT kondisi ( $M = 18,20, SD = 5,95$ ). Terakhir, terdapat interaksi yang signifikan antara kondisi dan tipe inferensi,  $F(3,177) = 16.07, p < .001, \eta^2 = .21$ . Sarana untuk interaksi ini ditunjukkan pada Tabel 7.3. Secara signifikan lebih banyak inferensi MP yang didukung dalam kondisi IT daripada kondisi OI,  $t(59) = 2.82, p = .007, d = 0.72$ . Sebaliknya, inferensi AC yang jauh lebih sedikit,  $t(59) = 5,58, p < .001, d = 1.43$ , dan inferensi MT  $t(59) = 2.09, p = .041, d = 0.53$ , didukung dalam kondisi TI. Ada juga sedikit lebih sedikit inferensi DA yang didukung dalam kondisi IT daripada kondisi OI,  $t(59) = 1.81, p = .075, d = 0.46$ . Data ini ditampilkan pada Gambar 6.1.

Untuk menyelidiki kemungkinan hipopenelitian konversi Evans (1977) untuk menjelaskan tingkat dukungan yang berbeda di seluruh kondisi, tingkat dukungan TI dibandingkan dengan tingkat dukungan OI yang dikonversi dengan cara yang sama (diringkas dalam Tabel 6.4). Ini memberikan 6,09 inferensi OI MP yang didukung dibandingkan dengan 3,43 inferensi AC IT, 3,97 inferensi OI DA yang didukung dibandingkan dengan 4,67 inferensi IT MT, 6,65 inferensi OI AC yang didukung dibandingkan dengan 7,20 inferensi IT MP, dan 5,74 inferensi OI MT yang didukung dibandingkan dengan 2,90 inferensi DA. Pertandingan yang disediakan di sini tidak sedekat dalam data Evans.



**Gambar 6.1:** Rata-rata tingkat pengesahan untuk setiap jenis inferensi dalam kondisi TI dan OI (bilah kesalahan mewakili  $\pm 1$  SE dari rata-rata).



**Tabel 6.4** Rata-rata tingkat dukungan untuk setiap inferensi TI dan inferensi OI yang dikonversi setara

Intervensi IT	Dukungan IT	Dukungan OI	Intervensi OI
MP	7.20	6.65	AC
DA	2.90	5.74	MT
AC	3.43	6.09	MP
MT	4.67	3.97	DA

## 6.5 INTERPRETASI KONDISIONAL

Selanjutnya, keempat indeks interpretasi dibandingkan antar kondisi. ANOVA 2x4 dengan satu faktor antara subjek: Tipe Kondisional (IT, OI) dan satu faktor dalam subjek: Interpretasi (MCI, DCI, BCI, CCI) mengungkapkan efek utama yang signifikan dari Interpretasi,  $F(3,177) = 5,067$ ,  $p = .002$ ,  $\eta^2 = .08$ , efek utama yang signifikan dari Tipe Kondisional,  $F(1,59) = 14,29$ ,  $p < .001$ ,  $\eta^2 = .20$ , dan interaksi signifikan,  $F(3,177) = 14,27$ ,  $p < .001$ ,  $\eta^2 = .20$  (lihat Gambar 7.2). Interaksi ditindaklanjuti dengan uji-t sampel independen untuk masing-masing dari empat interpretasi yang membandingkan rata-rata dalam kondisi TI dan OI.

MCI secara signifikan lebih tinggi pada kondisi IT ( $M=21.53$ ,  $SD=4.19$ ) dibandingkan kondisi OI ( $M=17.23$ ,  $SD=3.62$ ),  $t(59) = 4.30$ ,  $p < .001$ ,  $d = 1.10$ . DCI juga lebih tinggi pada kondisi IT ( $M=20.20$ ,  $SD=5.87$ ) dibandingkan pada kondisi OI ( $M=13.74$ ,  $SD=4.40$ ),  $t(59) = 4.87$ ,  $p < .001$ ,  $d = 1.25$ . Sebaliknya, BCI secara signifikan lebih rendah pada kondisi IT ( $M=18.20$ ,  $SD=5.95$ ) dibandingkan pada kondisi OI ( $M=22.45$ ,  $SD=5.27$ ),  $t(59) = 2.96$ ,  $p = .004$ ,  $d = 0,76$ . Tidak ada perbedaan antara CCI pada kondisi IT ( $M=19.07$ ,  $SD=3.27$ ) dan OI ( $M=19.03$ ,  $SD=3.07$ ),  $t(59) = .04$ ,  $p = .966$ ,  $d = 0,01$ .

## 6.6 DISKUSI

Dalam penelitian ini, mahasiswa matematika telah terbukti menjadi lebih cacat dalam penalaran kondisional abstrak (lihat Bab 4 dan 5). Di sini, sifat perilaku penalaran kondisional mahasiswa matematika diselidiki; apakah mereka menanggapi struktur logis yang mendasari masalah inferensi kondisional, seperti yang diprediksi oleh TFD, atau apakah mereka terpengaruh oleh ungkapan linguistik, seperti yang terjadi pada peserta non-matematika Evans (1977)?

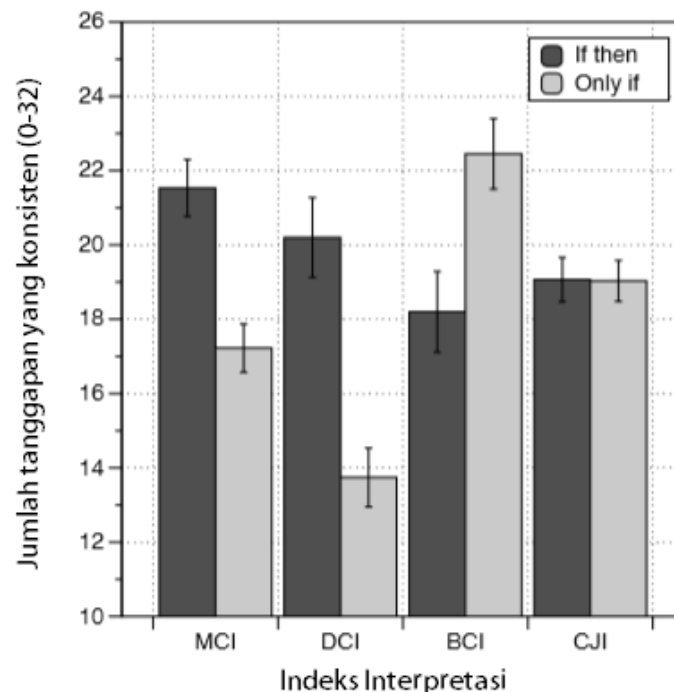
Sekelompok mahasiswa matematika sarjana tahun ketiga menyelesaikan tugas inferensi kondisional 32 item dalam salah satu dari dua kondisi: dengan frasa IT atau dengan frasa OI. Hasil ini sebagian besar sejalan dengan Evans (1977) yang mempelajari mahasiswa non-matematika. Peserta mendukung lebih banyak inferensi MP dalam kondisi IT, dan lebih banyak inferensi AC dan MT dalam kondisi OI. Ini menunjukkan dua hal: pertama bahwa mahasiswa matematika tidak memperlakukan pernyataan kondisional berdasarkan logika yang mendasarinya, dan kedua bahwa temuan Evans yang tidak terduga tentang tingkat dukungan yang lebih tinggi dari inferensi AC dalam kondisi OI daripada kondisi TI direplikasi.

Mahasiswa matematika tampaknya tidak menanggapi masalah inferensi kondisional berdasarkan struktur logis yang mendasarinya. Sebaliknya, mereka menafsirkan pernyataan kondisional secara berbeda tergantung pada ungkapan. Terdapat pengaruh utama kondisi baik untuk endorsement rate maupun indeks interpretasi, dengan penalaran mahasiswa lebih sesuai dengan materi dan interpretasi yang kurang tepat pada kondisi IT daripada pada kondisi

OI dan lebih sejalan dengan interpretasi bikondisional pada kondisi OI daripada pada kondisi OI kondisi TI.

Ini bertentangan dengan TFD, yang menyatakan bahwa "studi matematika tidak dapat digantikan oleh aktivitas lain apa pun yang akan melatih dan mengembangkan kemampuan logika murni manusia ke tingkat rasionalitas yang sama" (Oakley, 1949, hlm. 19), bahwa "melalui matematika kami juga ingin mengajarkan penalaran logis" (Amitsur dalam Sfard, 1998, hlm. 453), dan bahwa matematika "mendisiplinkan pikiran, mengembangkan penalaran logis dan kritis, dan mengembangkan keterampilan analitis dan pemecahan masalah ke tingkat yang tinggi (Smith, 2004, hlm. 11). Tampak jelas bahwa kutipan ini mendukung hipotesis penelitian mahasiswa matematika yang memegang pemahaman logika abstrak dan mampu menghindari pengaruh ungkapan atau konteks dengan 'pikiran disiplin' mereka. Sebaliknya, mereka sangat terpengaruh oleh ungkapan masalah, dengan kelompok TI dan OI merespons secara signifikan berbeda untuk tiga dari empat jenis inferensi (MP, AC, MT) dan sedikit berbeda untuk keempat (DA).

Ini tidak konsisten dengan TFD tetapi konsisten dengan hipotesis penelitian bahwa mahasiswa matematika belajar untuk mengasumsikan p dan alasan tentang q, yaitu menjadi lebih cacat dalam penalaran kondisional mereka. Meskipun ada kemungkinan bahwa mahasiswa menjadi lebih cacat dengan pernyataan OI dari waktu ke waktu, penelitian ini menunjukkan bahwa mereka tidak cacat dengan pernyataan OI seperti halnya dengan pernyataan TI pada tahun ketiga gelar sarjana matematika. Penjelasan yang diajukan untuk temuan ini, seperti dalam Bab 5 dan 6 adalah matematika menghadapkan mahasiswa pada pernyataan bentuk TI lebih sering daripada bentuk OI, dan bahwa keakraban dengan, praktik dengan, dan umpan balik pada masalah termasuk pernyataan TI mendorong mahasiswa kompetensi dengan mereka.



**Gambar 6.2:** Indeks interpretasi rata-rata dalam kondisi IT dan OI (bilah kesalahan mewakili  $\pm 1$  SE dari rata-rata).

Kurangnya paparan pernyataan OI dapat berarti bahwa mahasiswa matematika memperlakukan mereka dengan cara yang sama seperti yang dilakukan oleh non-matematika. Memang, Houston (2009) berpendapat bahwa sebagian besar pernyataan matematika berbentuk 'jika pernyataan A benar, maka pernyataan B benar', bahkan jika pernyataan tersebut sangat disamarkan (hal. 63). Lebih lanjut, ia berpendapat bahwa dalam matematika A dianggap benar, meskipun itu jelas tidak benar, dan kebenaran atau kesalahan B kemudian disimpulkan. Ini adalah deduksi bergerak maju, sedangkan 'A hanya jika B' dapat dilihat deduksi mundur di mana A dipertanyakan. Seperti yang ditunjukkan Houston, mungkin sulit untuk melihat ini sebagai setara dengan 'jika A maka B'. Kami kembali ke ide di bawah ini. Pertama, hipopenelitian konversi Evans (1977) dipertimbangkan, yang akan kita lihat menunjuk pada kesimpulan yang sama.

Data mendukung temuan Evans (1977) yang tidak terduga tentang lebih banyak inferensi AC yang didukung dalam kondisi OI daripada dalam kondisi TI. Evans menyarankan bahwa ini karena peserta melakukan konversi terlarang dari 'p hanya jika q' menjadi 'jika q maka p'. Dalam hal ini, inferensi OI AC setara dengan inferensi MP IT, yang akan menjelaskan tingkat dukungan yang tinggi. Namun, meskipun hipopenelitian konversi tampaknya memberikan kecocokan yang baik untuk data Evans dalam banyak kasus, tampaknya tidak begitu cocok untuk data ini. Ada perbedaan yang cukup besar antara tingkat dukungan MP dan MT dari masalah TI dan tingkat dukungan OI yang dikonversi setara.

**Tabel 6.5:** Persentase item yang didukung untuk setiap jenis inferensi di seluruh peserta dalam empat kelompok: grup IT matematika saat ini, grup OI matematika saat ini, grup non-matematika IT Evans (1977) dan grup non-matematika OI Evans.

	IT Matematika	IT Evan	OI Matematika	OI Evan
Modus Ponens	90.0%	100%	76.2%	87.5%
Penolakan pendahulunya	36.3%	68.8%	49.6%	50.0%
Penegasan akibat	42.9%	75.0%	83.1%	81.3%
Modus Tollens	58.3%	75.0%	71.8%	81.3%

Seperti yang disarankan di atas, ini mungkin terjadi karena mahasiswa matematika tampil mirip dengan non-matematika pada pernyataan OI tetapi dengan cara yang lebih material dan cacat daripada non-matematika pada pernyataan IT, sehingga kinerja dalam dua kondisi tidak dapat dibandingkan seolah-olah mereka adalah produk dari proses yang sama. Dengan kata lain, kinerja TI matematikawan secara kualitatif berbeda dengan kinerja OI mereka dan kinerja non-matematika pada kedua frasa. Pemeriksaan tingkat pengesahan pada Tabel 6.5 menunjukkan bahwa mahasiswa matematika dalam kelompok TI mendukung inferensi DA, AC, dan MT jauh lebih sedikit daripada kelompok TI non-matematika dalam studi Evans. Tingkat dukungan untuk kelompok OI matematika, bagaimanapun, tidak berbeda secara drastis dari kelompok OI non-matematika Evans.

Secara bersama-sama, perbedaan kinerja dalam kelompok TI dibandingkan dengan kelompok OI, dan kesamaan informal kinerja OI matematikawan dan non-matematika, menunjukkan bahwa keunggulan mahasiswa matematika dalam penalaran kondisional terbatas pada pernyataan TI, sedangkan kinerja mereka pada pernyataan OI mirip dengan non-matematikawan. Penjelasan untuk ini bisa jadi karena mahasiswa matematika menjadi

sangat terbiasa menilai pernyataan TI tetapi tidak terlalu banyak terpapar pernyataan OI (Houston, 2009). Berlawanan dengan TFD, ini membuat mereka berperilaku lebih cacat (dan secara material) pada pernyataan kondisional TI tetapi itu tidak mengubah interpretasi mereka terhadap persyaratan itu sendiri.

## **6.7 RINGKASAN DAN TEMUAN BARU**

1. Tampaknya keunggulan matematika dalam penalaran kondisional yang telah ditemukan sebelumnya bukan karena pemahaman yang lebih baik dari logika yang mendasarinya. Sebaliknya, mahasiswa matematika hanya tampak berhasil dengan pernyataan kondisional yang diutarakan 'jika  $p$  maka  $q$ ' dan tidak dengan ungkapan ' $p$  hanya jika  $q$ '.
2. Dalam kasus pernyataan OI, matematikawan mungkin berperilaku baik dalam banyak cara yang sama seperti non-matematika.
3. Apa artinya ini bagi TFD adalah klaim besar tentang matematika yang meningkatkan 'kemampuan logika murni' sebenarnya dilebih-lebihkan. Hubungan antara matematika dan keterampilan penalaran logis mungkin jauh lebih sempit dari yang diperkirakan sebelumnya.

## BAB 7

### PERAN HEURISTIK DALAM PENALARAN MAHASISWA

#### 7.1 PENDAHULUAN

Menurut model kognisi tri-proses Stanovich (2009a), penalaran terjadi melalui tiga tingkatan: tingkat heuristik, tingkat algoritmik dan tingkat reflektif. Pemrosesan heuristik, atau Tipe 1, cepat, otomatis, dan tidak menuntut sumber daya memori kerja. Pemrosesan tipe 2 disengaja dan menuntut memori kerja, diimplementasikan pada tingkat algoritmik dan didorong oleh tingkat reflektif (lihat Bagian 1.6). Ada kemungkinan bahwa peningkatan penalaran cacat mahasiswa matematika yang ditemukan di Bab 4 terjadi melalui satu atau lebih dari tiga level ini. Tujuan dari bab ini adalah untuk menyelidiki potensi tingkat heuristik menjadi sumber perubahan. Apakah kasus peningkatan kecenderungan mahasiswa matematika untuk penalaran yang cacat berasal dari tingkat kognisi heuristik?

Tingkat kognisi heuristik dicirikan oleh pemrosesan yang cepat dan otomatis yang sebagian besar tidak terlihat dari refleksi sadar. Output dari pemrosesan heuristik dan perasaan terkait kebenaran output tersedia untuk introspeksi, tetapi dapat dengan cepat dipengaruhi oleh refleksi Tipe 2. Ini berarti bahwa cara terbaik untuk mengukur respons tingkat heuristik adalah dengan akurasi yang dipercepat dan metode waktu reaksi. Dengan asumsi bahwa pemrosesan heuristik cepat dan otomatis dan pemrosesan Tipe 2 lebih lambat, dapat diasumsikan bahwa respons yang diberikan di bawah batas waktu yang singkat telah dihasilkan dengan sedikit atau tanpa input Tipe 2.

Beberapa penelitian sebelumnya telah menggunakan metode akurasi yang dipercepat untuk menyelidiki pemrosesan tingkat heuristik dalam tugas penalaran. Evans dan Curtis-Holmes (2005) memberi peserta 10 detik untuk menanggapi setiap item pada tugas silogisme bias keyakinan dan menemukan bahwa jumlah respons logis menurun sementara bias keyakinan meningkat, relatif terhadap kondisi waktu luang. Hal ini menunjukkan bahwa pengaruh keyakinan sebelumnya pada perilaku penalaran adalah proses heuristik dan pemrosesan Tipe 2 diperlukan untuk mengesampingkannya untuk memberikan jawaban yang logis secara normatif. Demikian pula ketika Gillard et al. (2009b) membatasi peserta mereka hingga 17 detik untuk menanggapi masalah yang menyebabkan heuristik proporsionalitas, jumlah tanggapan proporsional meningkat dan jumlah jawaban yang benar berkurang. Ini menunjukkan bahwa penalaran proporsional juga berbasis heuristik dan bahwa pemrosesan Tipe 2 diperlukan untuk menyimpannya dalam kasus di mana hal itu mengarah pada respons yang salah.

Saya menyadari hanya satu studi yang telah menyelidiki penalaran kondisional di bawah batasan waktu. Evans dkk. (2009) memberikan peserta mereka 48 masalah penalaran kondisional yang berbeda dalam kepercayaan dan validitas, dan mereka merespon baik dalam waktu 5 detik dari kesimpulan yang ditampilkan atau tanpa tekanan waktu. Masalah terdiri dari empat inferensi biasa, modus ponens (MP), penyangkalan antecedent (DA), penegasan konsekuensi (AC) dan modus tollens (MT), dan peserta menjawab dengan 'ya' atau 'tidak' untuk menunjukkan apakah mereka pikir kesimpulan harus mengikuti dari premis. Di bawah tekanan waktu peserta kurang cenderung untuk menerima kesimpulan secara keseluruhan, dan meskipun analisis jenis kesimpulan tidak dilaporkan, pemeriksaan sarana menunjukkan bahwa penurunan penerimaan tidak berbeda menurut jenis kesimpulan.

Mahasiswa matematika yang diselidiki dalam Bab 4 menjadi kurang cenderung untuk menerima kesimpulan MT, DA dan AC dari waktu ke waktu, sehingga menjadi lebih cacat dalam penalaran mereka. Tampaknya mungkin dari temuan Evans et al. (2009) bahwa ini bisa disebabkan oleh proses heuristik: jika ketergantungan yang lebih besar pada pemrosesan heuristik mengarah ke tingkat penolakan yang lebih tinggi dari inferensi kondisional, bisa jadi mahasiswa matematika menjadi lebih cenderung. mengandalkan intuisi mereka dari waktu ke waktu dan ini mengarah pada peningkatan tingkat penolakan mereka. Atau, bisa jadi kasus bahwa paparan pernyataan 'jika kemudian' mengarah pada perubahan tingkat heuristik di mana praktik dalam mengasumsikan  $p$  benar mengarah pada asumsi menjadi otomatis. Studi yang dilaporkan di bawah ini akan menyelidiki kemungkinan ini. Mahasiswa sarjana matematika dan sampel non-matematika menyelesaikan tugas Inferensi Kondisional standar dalam dua kondisi: satu di mana mereka dipaksa untuk merespons dengan cepat dan satu di mana mereka dapat menghabiskan waktu selama mereka suka bernalar. Beberapa hipotesis dapat diturunkan berdasarkan penelitian sebelumnya:

1. Para non-matematikawan akan menerima kesimpulan yang lebih sedikit secara keseluruhan dalam kondisi cepat daripada dalam kondisi lambat, sejalan dengan temuan Evans et al. (2009) dari sampel subjek-umum.
2. Mahasiswa matematika akan merespons lebih buruk dalam kondisi lambat daripada non-matematika, sejalan dengan temuan yang dilaporkan dalam Bab 5.
3. Jika tingkat heuristik bertanggung jawab atas cacat penalaran mahasiswa matematika, maka mahasiswa matematika akan tetap lebih cacat daripada mahasiswa non-matematika dalam kondisi cepat, dan tidak kurang cacat dalam kondisi cepat daripada dalam kondisi lambat .
4. Jika tingkat heuristik tidak bertanggung jawab atas cacat penalaran mahasiswa matematika, maka mereka akan merespon tidak lebih cacat daripada non-matematika dalam kondisi cepat, dan kurang cacat dalam kondisi cepat daripada dalam kondisi lambat.

Jika tingkat heuristik bertanggung jawab atas cacat penalaran mahasiswa matematika, dua penjelasan yang diusulkan di atas perlu dibedakan. Penjelasannya adalah mahasiswa matematika menjadi lebih bergantung pada output level heuristik, atau level heuristik mereka berubah dengan cara yang membuat mereka lebih fokus pada asumsi  $p$ . Hipotesis ini dapat dibedakan dengan membandingkan waktu reaksi kelompok (RT) dalam kondisi lambat: jika matematikawan lebih mengandalkan pemrosesan tingkat heuristik mereka, maka bahkan dalam kondisi lambat mereka harus merespons lebih cepat daripada non-matematika.

## **7.2 METODE**

### **7.2.1 Desain**

Mahasiswa matematika dan non-matematika menyelesaikan tugas Inferensi Kondisional Evans et al. (1995) dua kali: pertama dalam kondisi yang dipercepat dan kedua dengan waktu sebanyak yang mereka inginkan. Versi cepat selalu diselesaikan terlebih dahulu: dalam desain dalam mata pelajaran ada risiko bahwa menyelesaikan kondisi lambat terlebih dahulu dapat memungkinkan peserta untuk menghafal beberapa item atau mengingat jenis inferensi mana yang mereka anggap valid dan tidak valid, dan bahwa ini selanjutnya dapat memengaruhi mereka kinerja dalam kondisi cepat. Peserta juga menyelesaikan subset Matriks

Progresif Lanjutan Raven (RAPM) setelah Tugas Inferensi Kondisional untuk mengontrol perbedaan kecerdasan antar-kelompok.

### 7.2.2 Partisipan

Partisipan adalah 16 mahasiswa S1 dan S2 matematika serta 16 mahasiswa S1 dan S2 non-Matematika serta staf dari Stekom University. Ada 16 perempuan dan 16 laki-laki dan usia berkisar antara 18 sampai 51 ( $M=23.90$ ,  $SD=8.09$ ). Masing-masing dibayar Rp. 150.000 untuk waktu mereka.

### 7.2.4 Prosedur

Percobaan dilakukan pada komputer menggunakan E-prime 2.0. Pertama, peserta menyelesaikan studi yang tidak terkait yang melibatkan Tes Refleksi Kognitif. Untuk penelitian ini, tugas pertama yang mereka selesaikan adalah versi cepat dari tugas Inferensi Kondisional. Mereka melihat instruksi, 4 item latihan, dan 32 item nyata. Dalam setiap percobaan, premis kondisional disajikan sendiri selama 1,5 detik sebelum premis minor dan kesimpulan ditambahkan secara bersamaan selama 2,5 detik tambahan, di mana peserta diminta untuk merespons.

Batas waktu didasarkan pada uji coba yang menunjukkan bahwa pengaturan waktu ini mencegah peserta untuk dapat secara sadar merenungkan pertanyaan sambil mempertahankan respons tingkat heuristik (bias pencocokan memungkinkan item MP untuk siap diterima tanpa pemrosesan sadar, lihat lebih lanjut di bawah). Dalam versi lambat, peserta kembali melihat premis kondisional saja selama 1,5 detik, tetapi ketika premis minor dan kesimpulan ditambahkan, peserta tidak diizinkan untuk merespons selama 5 detik pertama, setelah itu mereka dapat mengambil selama mereka ingin merespons. Pembatasan ini mencegah versi cepat dari mempengaruhi peserta untuk merespons dengan cepat dalam versi lambat. Akhirnya, peserta menyelesaikan subset item dari RAPM (Stanovich & West, 1998), terdiri dari 18 item dengan batas waktu 10 menit. Peserta kemudian berterima kasih, dibayar, dan diberhentikan.

## 7.3 HASIL

### 7.3.1 Pemeriksaan manipulasi dan penilaian kovariat

Sebagai pemeriksaan manipulasi, tingkat endorsement pada item Modus Ponens dalam kondisi cepat dibandingkan dengan tingkat peluang dengan uji-t satu sampel. Karena MP dapat dengan mudah didukung melalui bias pencocokan, respons heuristik, kami berharap untuk menemukan tingkat dukungan yang tinggi bahkan ketika peserta dibatasi untuk pemrosesan tingkat heuristik, selama batas waktu tidak terlalu singkat. Batas waktu yang digunakan di sini (yang didasarkan pada uji coba) agak lebih pendek daripada yang digunakan dalam penelitian sebelumnya. Namun demikian, peserta mendukung kesimpulan MP pada tingkat peluang di atas ( $M=5,78$ ,  $SD=1,58$ , kemungkinan maksimum = 8),  $t(31) = 6,37, p < .001$ , menunjukkan bahwa batas waktu cukup bagi tingkat heuristik untuk menghasilkan sebuah tanggapan.

Tanggapan untuk tugas Inferensi Kondisional dikodekan ke dalam empat variabel untuk mencerminkan konsistensi dengan masing-masing dari empat interpretasi dari pernyataan kondisional: Indeks Kondisional Material, Indeks Kondisional Cacat, Indeks Bikondisional dan Indeks Konjungtif. Setiap variabel adalah skor dari 32, dengan skor yang lebih tinggi menunjukkan bahwa menanggapi lebih konsisten dengan interpretasi dari pernyataan kondisional.

Nilai RAPM secara signifikan lebih tinggi pada kelompok matematika ( $M=10.31$ ,  $SD=2.21$ ) dibandingkan kelompok non-matematika ( $M=6.19$ ,  $SD=2.10$ ),  $t(30) = 5.40$ ,  $p < .001$ ,  $d = 1.91$ , dan berkorelasi positif signifikan dengan Indeks Kondisional Material pada kondisi lambat,  $r(32) = .63$ ,  $p < .001$ , Indeks Kondisi Rusak pada kondisi lambat,  $r(32) = .59$ ,  $p < .001$ , dan berkorelasi negatif secara signifikan dengan Indeks Bikondisional dalam kondisi lambat,  $r(32) = .53$ ,  $p = .002$ . Respon RAPM dan Konjungtif dalam kondisi lambat tidak berkorelasi secara signifikan ( $p=.223$ ).

Respons untuk tugas Inferensi Kondisional dalam kondisi cepat tidak diharapkan berkorelasi dengan RAPM, karena kondisi cepat dirancang untuk mengurangi tingkat algoritme (diukur dengan RAPM) merespons. Ini memang kasus untuk Indeks Kondisional, Bikondisional, dan Konjungtif Cacat (semua  $ps > .130$ ). Korelasi yang signifikan antara RAPM dan Indeks Kondisi Rusak lambat dan korelasi tidak signifikan antara RAPM dan Indeks Kondisi Rusak cepat berbeda secara signifikan,  $t(29) = 2.17$ ,  $p = .019$ ,  $r = .48$ . Hal ini juga berlaku untuk korelasi signifikan antara RAPM dan Indeks Bikondisional lambat dan korelasi tidak signifikan antara RAPM dan Indeks Bikondisional cepat,  $t(29) = 2.04$ ,  $p = 0.025$ ,  $r = 0.46$ . Ini berfungsi sebagai pemeriksaan manipulasi kedua, menunjukkan bahwa batas waktu dalam kondisi cepat tidak terlalu lama untuk memungkinkan tingkat algoritmik mengganggu.

Namun, Material Conditional Index dalam kondisi cepat berkorelasi positif dengan RAPM,  $r(32) = .39$ ,  $p = .029$ . Ini mungkin karena fakta bahwa kelompok matematika memiliki skor kecerdasan yang lebih tinggi, dan seperti yang ditunjukkan pada bab-bab sebelumnya, lebih mungkin untuk merespon sesuai dengan kondisi materi daripada non-matematika. Penjelasan ini didukung oleh korelasi yang dilakukan pada masing-masing kelompok secara terpisah, yang menunjukkan tidak ada korelasi yang signifikan antara skor RAPM dan Indeks Kondisional Material dalam kondisi cepat pada kedua kelompok (keduanya  $ps > .400$ ).

## 7.4 ANALISIS UTAMA

### 7.4.1 *Hipopenelitian 1: Tingkat pengesahan pada kelompok non-matematika*

Evans dkk. (2009) menemukan bahwa sampel peserta non-matematika menerima lebih sedikit kesimpulan secara keseluruhan ketika dipaksa untuk merespons dengan cepat tugas Inferensi Kondisional. Untuk menyelidiki apakah ini juga terjadi di sini, tingkat dukungan kelompok nonmatematika dari masing-masing dari empat kesimpulan menjadi sasaran ANOVA  $2 \times 2$  dengan dua faktor dalam mata pelajaran: Waktu (cepat, lambat) dan Inferensi (MP, DA, AC, MT). Rata-rata tingkat endorsement pada setiap kondisi disajikan pada Gambar 8.1. Mahasiswa matematika tidak dimasukkan dalam analisis ini karena asumsi bahwa mereka bernalar dengan cara yang berbeda secara kualitatif dengan non-matematika (sebagaimana diuraikan dalam Bab 6).

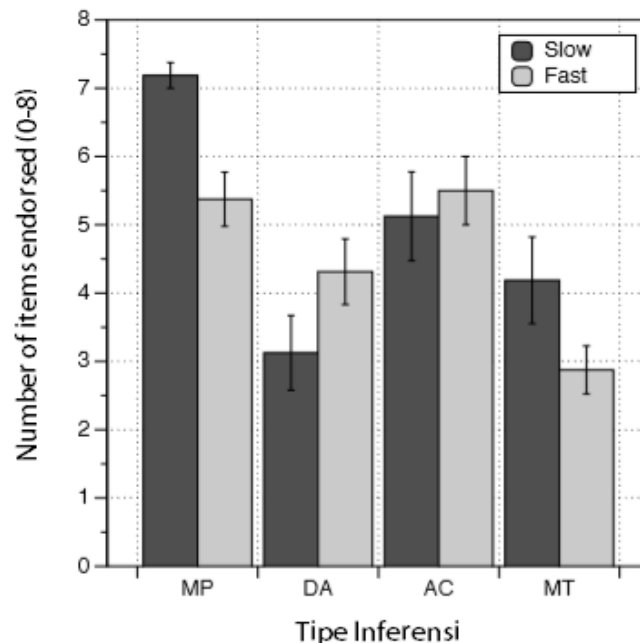
Berlawanan dengan temuan Evans et al (2009), tidak ada efek utama Waktu,  $F(1,45) = 1.41$ ,  $p = .254$ ,  $\eta^2 = .09$ , tetapi ada interaksi yang signifikan antara Waktu dan Inferensi,  $F(1,45) = 8.20$ ,  $p < .001$ ,  $\eta^2 = .35$ . Dalam kondisi cepat peserta menerima lebih sedikit inferensi MP,  $t(15) = 4.65$ ,  $p < .001$ ,  $d = 1.46$ , inferensi MT lebih sedikit,  $t(15) = 2.44$ ,  $p = .028$ ,  $d = 0.64$ , dan sedikit lebih banyak Inferensi DA,  $t(15) = 2.03$ ,  $p = .061$ ,  $d = 0.58$ , dibandingkan pada kondisi lambat. Sedangkan Evans dkk. (2009) menemukan penurunan keseluruhan dalam tingkat dukungan ketika peserta dipaksa untuk merespons dengan cepat, di sini tampak bahwa hanya kesimpulan yang valid yang secara signifikan lebih kecil kemungkinannya untuk didukung.



Ini diselidiki lebih lanjut dengan ANOVA  $2 \times 2$  dengan dua faktor dalam mata pelajaran: Waktu (cepat, lambat) dan Validitas (valid, tidak valid), yang mengungkapkan interaksi yang signifikan,  $F(1,15) = 24,94, p < .001, \eta^2 = 0,62$ . Uji-T dengan koreksi Bonferroni mengungkapkan bahwa peserta memang menerima lebih sedikit kesimpulan valid dalam kondisi cepat ( $M=8,25, SD=2,27$ ) dibandingkan dengan kondisi lambat ( $M=11,38, SD=2,75$ ),  $t(15) = 5,17, p < .001, d = 1.24$ , sedangkan tidak ada perbedaan yang signifikan antara jumlah inferensi tidak valid yang didukung pada kondisi cepat ( $M=8,25, SD=4,37$ ) dan lambat ( $M=9,81, SD=3,33$ ),  $t(15) = 1.61, p = .128, d = 0.40$ , lihat Gambar 7.2. Hipopenelitian 1, bahwa ahli matematika akan mendukung lebih sedikit kesimpulan dalam kondisi cepat daripada dalam kondisi lambat, oleh karena itu didukung sebagian.

#### 7.4.2 Hipopenelitian 2: Interpretasi kondisional tanpa tekanan waktu

Hipopenelitian 2 menyatakan bahwa mahasiswa matematika akan merespon lebih kurang baik daripada kelompok non-matematika dalam kondisi lambat, sesuai dengan hasil Bab 4. Indeks interpretasi rata-rata dalam kondisi lambat disajikan pada Gambar 7.3.

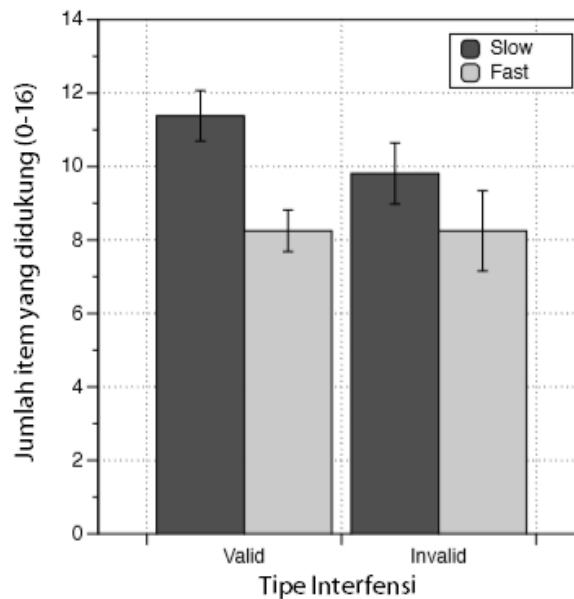


**Gambar 7.1:** Rata-rata tingkat pengesahan untuk setiap jenis inferensi (dari maksimum 8 item) dalam kondisi cepat dan lambat untuk kelompok non-matematika (bilah kesalahan mewakili  $\pm 1$  kesalahan standar rata-rata.).

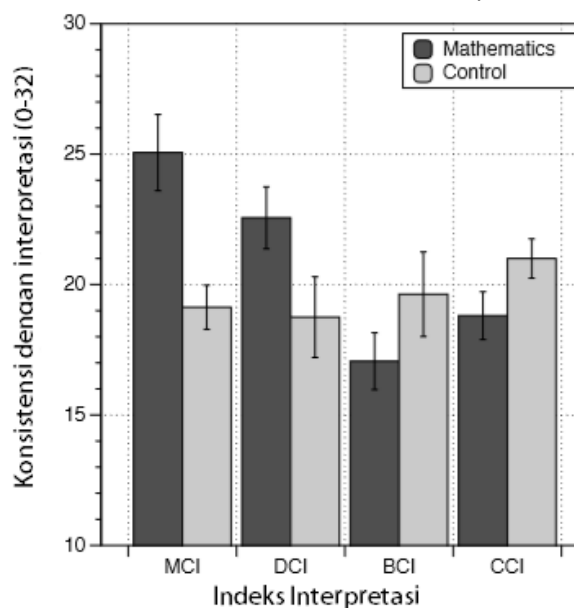
Indeks interpretasi dalam kondisi lambat dikenakan ANOVA  $2 \times 4$  dengan satu faktor dalam mata pelajaran: Interpretasi (Material, Cacat, Bikondisional, Konjungtif) dan satu faktor antar mata pelajaran: Kelompok (matematika, non-matematika). Karena indeks berasal dari rangkaian tanggapan yang sama dan oleh karena itu tidak independen, analisis ini diharapkan akan menunjukkan efek interpretasi yang utama. Sebenarnya ada efek utama Interpretasi yang sedikit signifikan,  $F(3,90) = 2.68, p = .052, \eta^2 = .082$ , dimana Material Conditional Index memiliki mean tertinggi ( $M=22.09, SD=5.06$ ), diikuti oleh Indeks Kondisional Cacat ( $M=20,66, SD=5,76$ ), kemudian Indeks Konjungtif ( $M=19,91, SD=3,48$ ), dan terakhir Indeks Bikondisional ( $M=18,34, SD=5,59$ ).

Ada juga interaksi yang signifikan antara Kelompok dan Interpretasi,  $F(3,90) = 5,02, p = 0,003, \eta^2 = 0,143$ . Uji-t yang direncanakan mengungkapkan bahwa kelompok matematika

memiliki Indeks Kondisional Material yang jauh lebih tinggi ( $M=25.06$ ,  $SD=5.85$ ) dibandingkan kelompok non-matematika ( $M=19.13$ ,  $SD=3.38$ ),  $t(30) = 3,52, p = .001$ ,  $d = 1,24$ . Kelompok matematika juga memiliki Indeks Kondisi Cacat yang sedikit lebih tinggi ( $M=22.56$ ,  $SD=1.19$ ) dibandingkan kelompok non-matematika ( $M=18.75$ ,  $SD=1.55$ ),  $t(30) = 1.96, p = .060$ ,  $d = 0,69$ . Analisis daya post hoc menghitung daya yang dicapai untuk analisis ini sebagai 0,61, menunjukkan bahwa ukuran sampel terlalu kecil meskipun ukuran efeknya cukup besar. Sebuah sampel dari 94 peserta akan diperlukan untuk menemukan efek yang signifikan dari ukuran ini dengan kekuatan 0,95. Kelompok matematika juga memiliki Indeks Konjungtif yang sedikit lebih rendah ( $M=18.81$ ,  $SD=3.66$ ) dibandingkan kelompok non-matematika ( $M=21.00$ ,  $SD=3.01$ ),  $t(30) = 1.85, p = .075$ ,  $d = 0,65$ . Tidak ada perbedaan antara kelompok dalam Indeks Bikondisional.



**Gambar 7.2:** Tingkat pengesahan untuk kesimpulan yang valid dan tidak valid dalam kondisi cepat dan lambat untuk kelompok non-matematika (bilah kesalahan menunjukkan  $\pm 1$  kesalahan standar rata-rata)



**Gambar 7.3:** Rata-rata indeks interpretasi pada masing-masing kelompok dalam kondisi lambat (bar error menunjukkan  $\pm 1$  standar error dari mean).

Meskipun pola hasil ini mencerminkan hasil Bab 4, interaksi antara Kelompok dan Interpretasi kehilangan signifikansi setelah RAPM dimasukkan sebagai kovariat,  $F(3,87) = 1,26, p = .293, \eta^2 = .04$ . Hipopenelitian 2 karena itu sebagian didukung; kelompok matematika memang memiliki Indeks Kondisional Cacat yang sedikit lebih tinggi daripada kelompok non-matematika ketika RAPM tidak diperhitungkan, tetapi mereka juga memiliki Indeks Kondisional Material yang jauh lebih tinggi daripada kelompok non-matematika, dan dengan ukuran efek hampir dua kali lebih besar ( $d = 1,24$  versus  $d = 0,69$ ). Selain itu, interaksi awal kehilangan signifikansi setelah skor RAPM diperhitungkan, menunjukkan bahwa perbedaan kelompok dalam interpretasi tidak dapat dipisahkan dari efek perbedaan dalam kemampuan kognitif umum. Namun, analisis ini memiliki kekuatan yang lebih rendah daripada analisis dalam studi tingkat AS.

Salah satu kemungkinan alasan tingginya tingkat MCI pada mahasiswa matematika dalam penelitian ini adalah karena mereka memiliki tingkat kecerdasan rata-rata yang lebih tinggi daripada mahasiswa dalam penelitian lain yang dilaporkan dalam penelitian ini. Dalam kelompok mahasiswa matematika saat ini, nilai rata-rata RAPM adalah 10,31 dengan batas waktu 10 menit, sedangkan dalam studi sarjana nilai rata-rata pada Waktu 1 adalah 11,10 dengan batas waktu 15 menit (kira-kira setara dengan 7,40 di bawah waktu 10 menit). membatasi). Pada mahasiswa tingkat AS, ada lagi batas waktu 15 menit dan nilai rata-rata RAPM pada Waktu 1 adalah 9,29 (kira-kira setara dengan 6,19 di bawah batas waktu 10 menit). Ini berarti menunjukkan bahwa kelompok mahasiswa matematika dalam studi saat ini mungkin memang lebih cerdas, dan ini pada gilirannya dapat menjelaskan skor MCI mereka yang tinggi.

**Tabel 7.1:** Nilai indeks rata-rata untuk setiap interpretasi pernyataan kondisional dalam kondisi cepat dan lambat untuk kelompok matematika dan non-matematika (kontrol) (standar deviasi dalam tanda kurung).

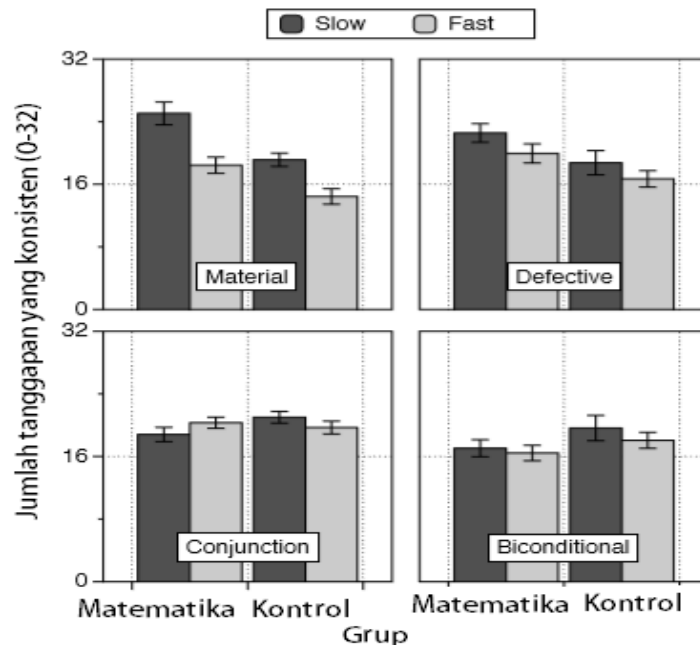
Grup	Waktu	Materi	Defektif	Bikondisional	Konjungtif
Matematika	Cepat	18.44 (4.11)	19.94 (4.91)	16.44 (3.95)	20.31 (2.82)
	Lambat	25.06 (5.85)	22.56 (4.75)	17.06 (4.36)	18.81 (3.66)
Kontrol	Cepat	14.44 (3.97)	16.69 (4.19)	18.06 (4.09)	19.69 (3.28)
	Lambat	19.13 (3.38)	18.75 (6.19)	19.63 (6.48)	21.00 (3.01)

#### 7.4.3 Hipopenelitian 3 dan 4: Interpretasi dari kondisi di bawah tekanan waktu

Hipopenelitian 3 menyatakan bahwa jika tingkat heuristik bertanggung jawab untuk penalaran yang lebih cacat kelompok matematika dibandingkan dengan kelompok non-matematika, maka perbedaan akan tetap pada kondisi cepat, dan Indeks Kondisi Cacat kelompok matematika tidak akan berbeda antara cepat dan lambat. kondisi. Hipopenelitian 4 alternatif menyatakan bahwa jika tingkat heuristik tidak bertanggung jawab, maka kelompok matematika akan merespon tidak lebih cacat daripada non-matematika dalam kondisi cepat, dan kurang cacat dalam kondisi cepat daripada dalam kondisi lambat.

Untuk membedakan antara hipopenelitian ini, skor inferensi kondisional menjadi sasaran ANOVA  $2 \times 4 \times 2$  dengan dua faktor dalam mata pelajaran: Waktu (lambat, cepat) dan Interpretasi (Material, Cacat, Bikondisional, Konjungtif), satu faktor antara mata pelajaran: Kelompok (matematika, non-matematika) dan satu kovariat: skor RAPM. Rerata indeks interpretasi masing-masing kelompok pada kondisi cepat dan lambat disajikan pada Tabel 8.1.

Seperti yang diharapkan, karena skor interpretasi tidak independen, ada efek utama interpretasi,  $F(3,87) = 3,38, p = .022, \eta^2 = .104$ . Anehnya, tidak ada efek utama Waktu,  $F < 1$ . Namun, ada interaksi tiga arah yang signifikan antara Grup, Waktu dan Interpretasi,  $F(3,87) = 2.89, p = .040, \eta^2 = .09$ , seperti yang digambarkan pada Gambar 8.4.



**Gambar 7.4:** Konsistensi dengan masing-masing dari empat interpretasi untuk matematika dan kelompok kontrol di bawah kondisi cepat dan lambat. Bilah kesalahan menunjukkan  $\pm 1$  s.e.m.

Interaksi ini diselidiki lebih lanjut dengan ANOVA 2x2 untuk masing-masing dari empat interpretasi, dengan satu faktor dalam mata pelajaran: Waktu (cepat, lambat), dan satu faktor antar mata pelajaran: Kelompok (matematika, non-matematika). RAPM juga dimasukkan sebagai kovariat dalam analisis tanggapan Material, Cacat dan Bikondisional, mengingat korelasi yang dilaporkan di atas. Untuk Indeks Kondisional Material, tidak ada interaksi yang signifikan antara Waktu dan Grup,  $F < 1$ . Untuk Indeks Kondisional Cacat, ada sedikit interaksi yang signifikan antara Waktu dan Grup,  $F(1,29) = 4,06, p = 0,053, \eta^2 = .12$ , seperti halnya untuk Indeks Bikondisional,  $F(1,29) = 3.22, p = .083, \eta^2 = .10$ , dan Indeks Konjungtif,  $F(1,30) = 3.91, p = .057, \eta^2 = .12$ .

Mengingat bahwa Indeks Kondisi Cacat adalah yang paling menarik di sini (ini meningkat bersamaan dengan studi matematika di mahasiswa tingkat AS di Bab 5) dan menghasilkan nilai-p terkecil ( $p=.053$ ) dan ukuran efek terbesar ( $\eta^2 = .12$ ) dari interaksi ANOVA 2x2 yang dilaporkan di atas, sifat interaksi yang sedikit signifikan diselidiki dengan uji-t. Untuk Indeks Kondisi Cacat, nilai kelompok matematika secara signifikan lebih tinggi pada kondisi lambat ( $M=22.56, SD=4.75$ ) dibandingkan pada kondisi cepat ( $M=19.94, SD=4.91$ ),  $t(16) = 2.49, p = 0,025, d = 0,54$ , sedangkan nilai kelompok nonmatematika tidak berbeda berdasarkan Waktu ( $p=.21$ ). Indeks Kondisional Cacat kelompok matematika tetap sedikit lebih tinggi daripada kelompok nonmatematika di bawah kedua cepat,  $t(30) = 2.02, p = .053, d = 0.71$ , dan lambat,  $t(30) = 1.96, p = .060, d = 0,69$ , kondisi. Hipopenelitian 3 dan 4 diusulkan

sebagai alternatif satu sama lain, namun tampaknya keduanya tidak dapat ditolak secara meyakinkan.

Akhirnya, RT masing-masing kelompok dalam kondisi lambat dibandingkan untuk menguji hipotesis bahwa mahasiswa matematika mungkin lebih mengandalkan pemrosesan Tipe 1 daripada non-matematika. Uji-t sampel independen menunjukkan tidak ada perbedaan yang signifikan antar kelompok di RT,  $t(29) = .41, p = .685, d = 0.15$ .

## 7.5 DISKUSI

Tujuan dari Bab ini adalah untuk menyelidiki apakah tingkat kognisi heuristik, seperti yang dijelaskan oleh Stanovich (2009a), berpotensi menjadi sumber perubahan yang ditemukan dalam perilaku penalaran mahasiswa matematika yang dijelaskan dalam Bab 5. Mahasiswa matematika pada tingkat A ditemukan menjadi lebih cacat dalam penalaran mereka dari waktu ke waktu. Hal ini ditandai dengan pengurangan dukungan mereka terhadap inferensi DA, AC dan MT. Evans dkk. (2009) menemukan bahwa ketika peserta dipaksa untuk mengandalkan pemrosesan tingkat heuristik, dukungan mereka terhadap keempat kesimpulan menurun. Oleh karena itu tampak masuk akal bahwa ketergantungan yang lebih besar pada tingkat heuristik bisa menjadi sumber alasan cacat matematikawan.

Sekelompok mahasiswa sarjana dan pascasarjana matematika dan sekelompok sarjana dan staf non-matematika di Universitas menyelesaikan tugas Inferensi Kondisional dua kali: pertama dengan total 4 detik untuk membaca dan menanggapi setiap item, dan kedua dengan minimal 6,5 detik untuk membaca masalah sebelum diizinkan untuk menanggapi. Batas waktu cepat terbukti cukup lama untuk tingkat heuristik untuk memproses masalah (ditunjukkan dengan penerimaan inferensi MP di atas peluang) dan cukup pendek untuk mencegah tingkat algoritmik mempengaruhi respons (ditunjukkan oleh kurangnya korelasi antara RAPM dan menanggapi tugas Inferensi Kondisional dalam kondisi cepat).

Hipotesis 1 menyatakan bahwa non-matematikawan akan mendukung lebih sedikit inferensi dalam kondisi cepat daripada dalam kondisi lambat, seperti yang terjadi pada peserta Evans et al. (2009). Ini sebagian didukung: daripada penurunan umum dalam tingkat dukungan, hanya item yang valid yang didukung lebih sedikit dalam kondisi cepat. Ini menunjukkan bahwa pemrosesan Tipe 2 bertanggung jawab untuk mendukung inferensi MP dan MT dengan benar ketika non-matematikawan beralasan tanpa tekanan waktu, sedangkan tingkat heuristik bertanggung jawab untuk mendukung inferensi DA dan AC yang salah. Hal ini juga didukung oleh dukungan inferensi DA yang sedikit lebih tinggi dalam kondisi cepat.

Hipotesis 2 menyatakan bahwa mahasiswa matematika akan merespon lebih cacat daripada non-matematika dalam kondisi lambat. Indeks Kondisi Cacat Mahasiswa matematika sedikit lebih tinggi daripada non-matematikawan, dan ukuran efeknya sedang-besar ( $d = 0,69$ ). Namun, perbedaan ini dijelaskan oleh skor RAPM kelompok matematika yang lebih tinggi. Hal yang sama berlaku untuk Material Conditional Index, yang secara signifikan lebih tinggi pada kelompok matematika daripada kelompok non-matematika dengan ukuran efek yang besar ( $d = 1,24$ ), meskipun hanya sebelum RAPM telah diperhitungkan. Perbandingan skor RAPM dalam kelompok matematika dari penelitian ini dan dari studi tingkat sarjana dan AS menunjukkan bahwa mahasiswa matematika yang diselidiki di sini memiliki tingkat kecerdasan rata-rata yang lebih tinggi, yang dapat menjelaskan tingkat kesesuaian mereka yang tinggi terhadap persyaratan materi.

Dalam Bab 4, mahasiswa matematika tingkat AS ditemukan menjadi semakin cacat dalam penalaran mereka dari waktu ke waktu dibandingkan dengan kelompok kontrol, dan pada tingkat lebih rendah mereka lebih sesuai dengan interpretasi Material Conditional dari waktu ke waktu. Di sini, tren ini tampaknya telah berubah. Mungkin terjadi bahwa pada awal studi matematika pasca-wajib (yaitu tingkat AS) mahasiswa menjadi kurang bikondisional dan lebih cacat dalam penalaran mereka. Dengan kata lain, mereka mulai mendukung lebih sedikit inferensi DA, AC, dan MT. Kemudian dalam studi matematika mereka (yaitu di tingkat sarjana), mungkin mahasiswa bergerak lebih ke arah interpretasi materi kondisional, di mana mereka masih mendukung lebih sedikit inferensi DA dan AC, tetapi kembali ke menerima lebih banyak inferensi MT dengan benar.

Sejalan dengan saran ini, data yang disajikan dalam Bab 5 menunjukkan bahwa mahasiswa matematika tahun pertama menjadi lebih banyak materi dalam interpretasi mereka terhadap kondisi dari waktu ke waktu, meskipun interaksi antara Waktu dan Grup tidak signifikan. Selanjutnya, mahasiswa matematika sarjana tahun ketiga di Bab 6 memiliki MCI yang sedikit lebih tinggi daripada DCI dalam kondisi 'jika maka' (lihat Gambar 6.2). Sebagai alternatif, kontradiksi antara mahasiswa AS di Bab 5 dan mahasiswa pascasarjana di sini dapat disebabkan oleh efek seleksi, di mana (secara keseluruhan) hanya mahasiswa tingkat A paling cerdas yang melanjutkan studi di tingkat pascasarjana. Hipopenelitian-hipopenelitian tersebut sebaiknya diuji dengan studi longitudinal yang durasinya jauh lebih lama daripada yang disajikan dalam penelitian ini, misalnya dari awal jenjang AS hingga akhir jenjang sarjana, atau bahkan hingga studi pascasarjana.

Hipopenelitian 3 menyatakan bahwa jika tingkat heuristik bertanggung jawab untuk penalaran yang lebih cacat kelompok matematika dibandingkan dengan kelompok non-matematika, maka perbedaan akan tetap pada kondisi cepat, dan Indeks Kondisi Cacat kelompok matematika tidak akan berbeda antara cepat dan lambat. kondisi. Hipopenelitian 4 alternatif menyatakan bahwa jika tingkat heuristik tidak bertanggung jawab, maka kelompok matematika akan merespon tidak lebih cacat daripada non-matematika dalam kondisi cepat, dan kurang cacat dalam kondisi cepat daripada dalam kondisi lambat.

Tak satu pun dari hipopenelitian ini dapat dikesampingkan secara meyakinkan berdasarkan data yang disajikan di sini. Di satu sisi, mahasiswa matematika merespons kurang cacat di bawah tekanan waktu daripada yang mereka lakukan tanpa tekanan waktu, menunjukkan bahwa tingkat algoritmik setidaknya sebagian bertanggung jawab atas kecenderungan ini, mendukung Hipopenelitian 4. Di sisi lain, mahasiswa matematika merespons sedikit lebih banyak. cacat daripada kelompok non-matematika di bawah kondisi cepat maupun di bawah kondisi lambat, menunjukkan bahwa tingkat heuristik di beberapa bagian bertanggung jawab atas kecenderungan mereka untuk merespon secara cacat sesuai dengan Hipopenelitian 3. Namun, kelompok matematika hanya sedikit lebih cacat daripada kelompok non-matematika di bawah kedua batas waktu, dan ini mungkin menunjukkan bahwa Hipopenelitian 4 memiliki lebih banyak dukungan: perbedaan antara Indeks Kondisional Cacat mahasiswa matematika di bawah kondisi cepat dan lambat secara signifikan berbeda dengan ukuran efek  $d = 0,54$ , menunjukkan bahwa respons yang rusak adalah setengah dari standar deviasi yang lebih tinggi dalam kondisi lambat. Ini dapat mencerminkan pengaruh pemrosesan Tipe 2 dalam kondisi lambat, baik yang berasal dari tingkat algoritmik atau reflektif.

Mungkin hasil yang beragam menunjukkan bahwa tingkat heuristik sebagian, tetapi tidak seluruhnya, bertanggung jawab atas perilaku penalaran mahasiswa matematika yang cacat. Kecenderungan untuk menganggap kasus-kasus tidak-p tidak relevan (merek dari penalaran kondisional yang rusak) dapat dimulai sebagai proses sadar dan secara bertahap menjadi tertanam ke dalam tingkat heuristik, mirip dengan keterampilan mengemudi (Lewin, 1982; Newstead, 2000). Bisa jadi paparan berulang terhadap inferensi ke depan (seperti yang dibahas dalam Bab 4) menciptakan kebiasaan bawah sadar untuk menganggap kasus bukan-p tidak relevan. Kecenderungan ini dapat diperkuat oleh penalaran Tipe 2 melalui proses pembenaran seperti yang dijelaskan oleh Evans (2006, 2011), yang pada gilirannya dapat menumbuhkan kecenderungan cacat dalam masalah-masalah berikutnya, yang mengarah ke Indeks Kondisional Cacat yang lebih tinggi dalam kondisi lambat. Jika demikian halnya, bisa jadi refleksi sadar yang meningkat pada akhirnya memicu mahasiswa matematika untuk lebih bergerak ke arah respon kondisional materi di kemudian hari dalam pendidikan matematika mereka.

Disarankan berdasarkan temuan Evans et al. (2009) bahwa penalaran cacat mahasiswa matematika dapat disebabkan oleh ketergantungan yang lebih besar pada pemrosesan tingkat heuristik sebagai lawan dari perubahan sifat pemrosesan tingkat heuristik. Jika ini masalahnya, kami berharap ahli matematika merespons lebih cepat daripada non-matematika bahkan tanpa tekanan waktu. Ini tidak terjadi: tidak ada perbedaan yang signifikan antara RT rata-rata kedua kelompok dalam kondisi lambat. Hal ini konsisten dengan hipotesis bahwa setiap perbedaan tingkat heuristik antar kelompok disebabkan oleh sifat pemrosesan daripada memilih Tipe 1 daripada pemrosesan Tipe 2.

## 7.6 RINGKASAN DAN PENEMUAN BARU

1. Tingkat heuristik tidak sepenuhnya bertanggung jawab untuk penalaran cacat sarjana matematika dibandingkan dengan kelompok kontrol: mahasiswa matematika merespons secara signifikan kurang cacat ketika dibatasi pada pemrosesan tingkat heuristik.
2. Tingkat heuristik juga tidak sepenuhnya berlebihan: mahasiswa matematika sedikit lebih cacat daripada kelompok kontrol bahkan ketika terbatas pada pemrosesan tingkat heuristik.
3. Tampaknya pemrosesan Tipe 1 dan Tipe 2 memainkan peran dalam perilaku penalaran kondisional mahasiswa matematika.

Dalam beberapa hal, pemrosesan Tipe 1 dan Tipe 2 memainkan peran dalam perilaku penalaran mahasiswa matematika; tingkat heuristik terus bekerja untuk mengarahkan perhatian kita ke aspek yang relevan dari lingkungan kita, dan ketika waktu tidak dibatasi pemrosesan Tipe 2 juga selalu terlibat sampai batas tertentu, bahkan jika hanya untuk menyetujui respons yang dihasilkan secara heuristik (Evans, 2011).

Bab berikutnya menyelidiki potensi fungsi eksekutif, yaitu, efisiensi pemrosesan level algoritmik, untuk bertanggung jawab atas perbedaan antara perilaku penalaran mahasiswa matematika dan non-matematika. Sekelompok keterampilan fungsi eksekutif mahasiswa non-matematika diukur, bersama dengan perilaku Inferensi Kondisional mereka, untuk melihat apakah interpretasi yang berbeda dari pernyataan kondisional dikaitkan dengan memori kerja, penghambatan, atau keterampilan pergeseran yang lebih baik atau lebih buruk.

## BAB 8

### PERAN FUNGSI EKSEKUTIF DALAM BERNALAR KONDISIONAL

#### 8.1 PENDAHULUAN

Telah ditetapkan dalam Bab 4 dan (pada tingkat lebih rendah) bab 5 bahwa mempelajari matematika pada tingkat lanjutan dikaitkan dengan perubahan dalam keterampilan penalaran logis, seperti yang disarankan TFD. Namun, alih-alih bernalar secara lebih normatif, ditemukan bahwa peserta yang belajar matematika semakin menolak inferensi DA, AC dan MT dan semakin menerima inferensi MP, membuat penalaran mereka semakin rusak. Pertanyaan yang tersisa adalah bagaimana mekanisme perubahan ini. Dalam model kognisi tri-proses Bab 1 Stanovich (2009a) diidentifikasi sebagai titik awal yang berguna untuk mempersempit kemungkinan. Model tri-proses mengusulkan bahwa kognisi terjadi melalui tiga tingkatan: tingkat heuristik, tingkat algoritmik dan tingkat reflektif. Fokus bab ini adalah pada tingkat algoritmik: kapasitas komputasi dan efisiensi yang tersedia untuk penalaran yang sadar dan penuh upaya.

Dikatakan dalam tinjauan literatur bahwa kecerdasan umum dan fungsi eksekutif adalah konstruksi yang membentuk bagian dari tingkat algoritmik. Studi longitudinal yang dilaporkan dalam Bab 4 menyarankan bahwa kecerdasan bukanlah mekanisme peningkatan penalaran yang rusak, tetapi tujuan dari bab ini adalah untuk menilai potensi fungsi eksekutif menjadi sebuah mekanisme.

Fungsi eksekutif mengatur bagaimana kita menggunakan sumber daya kognitif kita untuk menyelesaikan tugas. Diperkirakan ada tiga fungsi eksekutif: memori kerja, penghambatan, dan pergeseran. Memori kerja adalah kemampuan untuk menyimpan dan memperbarui informasi dalam pikiran sadar. Individu bervariasi dalam seberapa banyak informasi yang dapat mereka ingat, seberapa akurat mereka dapat memantau informasi ini dan seberapa efektif mereka dapat menghapus informasi yang tidak lagi diperlukan dan menggantinya dengan informasi baru.

Penghambatan adalah kemampuan untuk menahan diri dari membuat tanggapan yang tidak diinginkan dan dapat berupa fisik atau kognitif. Contoh penghambatan fisik adalah tidak melihat gangguan ketika diminta untuk memfokuskan visi Anda pada titik yang ditentukan. Contoh penghambatan kognitif adalah membaca nomor telepon baru Anda segera setelah mengubahnya dari yang lama. Pergeseran mengacu pada kemampuan untuk mengalihkan perhatian antara tugas-tugas berbeda yang dilakukan secara bersamaan, misalnya, mengganti CD dengan aman saat mengemudi, atau memasak sambil menonton film tanpa kehilangan jejak tugas mana pun.

Tiga fungsi eksekutif telah terbukti secara jelas dapat dipisahkan satu sama lain dan dari kecerdasan. Studi yang dilaporkan di sini menyelidiki kontribusi masing-masing fungsi eksekutif terhadap perilaku penalaran kondisional dalam sekelompok mahasiswa non-matematika. Jika ada hubungan yang ditemukan, itu akan menunjukkan mekanisme potensial yang melaluinya mempelajari matematika dapat mengubah perilaku penalaran. Secara khusus, mungkin kasus belajar matematika meningkatkan memori kerja seseorang, penghambatan atau kemampuan bergeser, dan ini pada gilirannya menyebabkan kemampuan penalaran berubah. Agar hal ini terjadi, perbedaan individu dalam fungsi eksekutif perlu dikaitkan



dengan perbedaan individu dalam perilaku penalaran. Studi yang dilaporkan di sini tidak akan menguji apakah fungsi eksekutif adalah mekanisme di balik TFD, hanya apakah mungkin memang demikian.

Penelitian sebelumnya telah menunjukkan bahwa kapasitas memori kerja terlibat dalam berbagai bentuk penalaran, termasuk penalaran kondisional. Penelitian awal menyelidiki penalaran kondisional dari pandangan model memori kerja Baddeley & Hitch (1974, 1986). Menurut model, memori kerja terdiri dari tiga komponen: eksekutif pusat, loop fonologis dan sketsa visuo-spasial. Dua komponen terakhir adalah sistem penyimpanan jangka pendek untuk informasi verbal dan visuo-spasial, sementara eksekutif pusat mengontrol perhatian dan mengatur arus informasi ke dan dari sistem budak. Toms, Morris dan Ward (1993) menemukan bahwa hanya komponen eksekutif pusat yang direkrut untuk penalaran kondisional, sedangkan sketsa pad visuo-spasial dan lingkaran fonologis tidak berhubungan. Mereka menyarankan bahwa penalaran kondisional membutuhkan media memori kerja abstrak sebagai lawan dari verbal atau visuo-spasial.

Temuan hubungan antara penalaran kondisional dan memori kerja telah direplikasi secara konseptual dalam berbagai penelitian. De Neys, Schaeken dan d'Ydewalle (2005) menemukan bahwa peserta dengan rentang memori kerja yang tinggi beralasan lebih sesuai dengan kondisi materi daripada peserta dengan rentang memori kerja yang rendah. Selanjutnya, Verschueren et al. (2005) menemukan bahwa kapasitas memori kerja tidak hanya terkait dengan kinerja penalaran, tetapi juga menentukan apakah seorang penalaran akan menggunakan strategi probabilistik (kapasitas rendah) atau kontracontoh (kapasitas tinggi) untuk memecahkan masalah. Namun, hubungannya mungkin tidak langsung: Handley, Capon, Copp dan Harper (2002) menemukan bahwa kinerja pada tugas penalaran kondisional hanya terkait dengan memori kerja verbal, bukan memori kerja spasial. Tampaknya dari temuan mereka seolah-olah penalaran kondisional yang berhasil bergantung pada pemrosesan simultan dan penyimpanan representasi verbal, yang bertentangan dengan temuan Toms et al. (1993).

Meskipun telah diketahui dengan baik bahwa memori kerja dalam beberapa hal penting untuk penalaran, penelitian tentang aspek penghambatan dan pergeseran fungsi eksekutif dan hubungannya dengan kemampuan penalaran lebih jarang. Sejauh pengetahuan saya, hubungan antara penalaran kondisional dan kemampuan pergeseran belum diselidiki dan beberapa penelitian telah menyelidiki peran penghambatan dalam penalaran. Dalam sebuah studi tentang kemampuan penalaran anak-anak berusia 10 tahun, Handley et al. (2004) menemukan bahwa sementara penalaran dengan masalah berbasis keyakinan terkait dengan memori kerja dan keterampilan penghambatan, penalaran netral keyakinan hanya terkait dengan memori kerja. Demikian pula, Markovits dan Doyon (2004) menemukan bahwa kemampuan untuk menghambat informasi yang tidak relevan terkait dengan keberhasilan dalam memecahkan masalah penalaran kondisional tematik, serta, dan secara terpisah dari, memori kerja.

Penghambatan, kemudian, mungkin hanya menjadi alat kognitif yang diperlukan ketika seorang penalaran perlu mendekontekstualisasikan masalah untuk menyelesaikannya, yaitu ketika mereka perlu menghambat keyakinan mereka sebelumnya. Jika ini masalahnya, maka tugas Inferensi Kondisional abstrak yang digunakan di seluruh penelitian ini, dan dalam penelitian yang dilaporkan dalam bab ini, mungkin tidak memerlukan keterampilan penghambatan, tetapi tampaknya kinerja akan terkait dengan memori kerja.

Bergerak di luar kinerja tugas Inferensi Kondisional secara keseluruhan, studi yang disajikan di sini menawarkan kesempatan untuk menyelidiki Kesimpulan Negatif dan bias Premis Afirmatif, dan beberapa hipopenelitian dapat dirumuskan. Negatif Kesimpulan Bias (NCB) adalah kecenderungan untuk menarik lebih banyak kesimpulan dengan kesimpulan negatif daripada dengan kesimpulan positif dan paling sering diamati pada kesimpulan DA dan MT (Schroyens et al., 2001). NCB telah dijelaskan dalam dua cara yang dapat diuji di sini. Satu saran adalah NCB adalah bias heuristik, di mana para pemikir berasumsi bahwa 'bukan p' lebih mungkin benar daripada 'p' (ada lebih banyak hal non-manusia daripada manusia, dan lebih banyak non-tabel daripada ada tabel, misalnya) dan lebih bersedia menerimanya. Atau, NCB mungkin menjadi bias pada tingkat pemrosesan Tipe 2. Evans dkk. (1995) telah mengusulkan bahwa masalahnya terletak pada negasi ganda yang melekat pada inferensi DA dan MT dengan kesimpulan afirmatif. Saat membuat inferensi MT dari kondisi 'Jika A maka 3', premis 'bukan 3' mengarah ke kesimpulan negatif 'bukan A'. Namun, ketika membuat inferensi yang sama dari kondisi 'Jika bukan A maka 3', langkah tambahan dilibatkan untuk mencapai kesimpulan afirmatif yang diperlukan: 'bukan 3' menyiratkan 'bukan (bukan A)', yang perlu dikonversi menjadi 'A'.

Evans dkk. (1995) berpendapat bahwa penalaran tidak mudah melihat kesetaraan 'tidak (bukan p)' dan 'p', dan ini adalah sumber kesulitan. Hubungan kuadrat yang ditemukan antara NCI dan DCI di Bab 5 menunjukkan bahwa penjelasan terakhir lebih mungkin terjadi - NCI tampaknya merupakan produk sampingan dari penalaran sistematis daripada proses heuristik.

Kedua hipopenelitian bersaing membuat prediksi yang berbeda untuk hubungan antara NCB dan fungsi eksekutif. Jika NCB berasal dari proses tingkat heuristik, maka penghambatan yang lebih baik mungkin terkait dengan keberhasilan yang lebih besar dalam menghindarinya. Jika, di sisi lain, NCB berasal dari perjuangan untuk menyelesaikan langkah ekstra logis yang diperlukan untuk membuat inferensi MT dan DA dengan kesimpulan negatif, maka kita mungkin mengharapkan memori kerja yang lebih baik untuk dikaitkan dengan tingkat bias yang lebih rendah.

Bias Premis Afirmatif (APB) adalah kecenderungan untuk menerima lebih banyak kesimpulan dari premis afirmatif daripada dari premis negatif, terutama ketika premis negatif tersirat. Misalnya, inferensi 'jika tidak-A maka 3; SEBUAH; oleh karena itu not-3' diterima lebih sering daripada inferensi 'jika A maka 3; D; oleh karena itu bukan-3', meskipun keduanya merupakan inferensi DA yang tidak valid. Ini mungkin karena bias yang cocok dalam pemrosesan Tipe 1, di mana premis 'A' lebih jelas terkait dengan kondisional daripada premis 'D' (Evans & Handley, 1999). Seperti NCB, jika APB memang disebabkan oleh kesalahan pemrosesan Tipe 1, maka kemampuan untuk menghindarinya mungkin terkait dengan penghambatan.

Untuk meringkas, ada beberapa prediksi untuk studi saat ini:

1. Tiga komponen fungsi eksekutif tidak akan berhubungan.
2. Memori kerja akan berkorelasi positif dengan MCI.
3. Mengingat sifat tugas yang abstrak, penghambatan tidak akan dikaitkan dengan perilaku penalaran.
4. NCB akan terkait dengan penghambatan ATAU skor memori kerja.
5. APB akan berhubungan dengan skor inhibisi.

Tidak ada saran dari penelitian sebelumnya bahwa skor yang bergeser akan (atau tidak akan) terkait dengan kinerja pada tugas Inferensi Kondisional yang abstrak, jadi aspek ini bersifat eksplorasi.

## 8.2 METODE

### 8.2.1 Partisipan

Sembilan puluh empat mahasiswa pascasarjana dari berbagai mata kuliah teknik mengambil bagian tanpa dibayar selama sesi lab untuk modul statistik pengantar tahun kedua. Mereka kemudian menganalisis data untuk tugas kuliah mereka. Semua peserta memberikan persetujuan untuk data mereka yang akan digunakan untuk tujuan penelitian.

### 8.3 PENGUKURAN

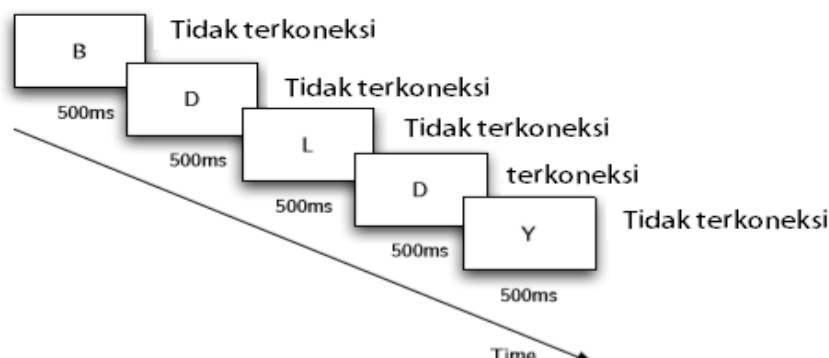
Ada empat ukuran, satu untuk masing-masing memori kerja, penghambatan, pergeseran, dan penalaran kondisional. Semua tugas diprogram dan dikelola dalam E-Prime versi 2.0.

#### 8.3.1 Memori kerja

Memori kerja diukur menggunakan tugas 2-kembali. Peserta melihat serangkaian huruf yang disajikan secara berurutan dan diinstruksikan untuk menekan tombol 'Target' ketika huruf di layar cocok dengan yang disajikan dua huruf kembali. Untuk semua huruf lainnya, mereka menekan tombol 'Bukan target'. Contoh urutan percobaan ditunjukkan pada Gambar 9.1. Ada 90 percobaan, 30 di antaranya adalah target, dan tugas itu didahului dengan sesi latihan 20 percobaan. Huruf disajikan selama 500 ms dengan jarak antar huruf 1000 ms. Ukuran yang diambil adalah akurasi rata-rata.

#### 8.3.2 Inhibisi

Sebuah versi tugas Stroop (Stroop, 1935) digunakan sebagai ukuran penghambatan kognitif. Dalam kondisi tidak ada konflik, peserta melihat rangkaian simbol @ yang disajikan dalam font berwarna dan diinstruksikan untuk mengidentifikasi warna font tersebut. Dalam kondisi konflik, peserta melihat nama warna disajikan dalam font berwarna berbeda, misalnya, kata 'biru' disajikan dalam font merah, dan kembali diperintahkan untuk mengidentifikasi warna font, mengabaikan kata yang bertentangan. Kondisi diblokir dan masing-masing terdiri dari 40 percobaan yang didahului dengan 10 percobaan latihan. Lima warna digunakan - merah, biru, hijau, kuning dan ungu. Peserta menanggapi setiap percobaan dengan menekan salah satu dari lima tombol yang diidentifikasi dengan stiker berwarna. Partisipan diinstruksikan untuk merespon secepat dan seakurat mungkin dan stimulus ditampilkan sampai respon.



**Gambar 8.1:** Contoh urutan percobaan dari tugas 2-kembali

Kemampuan penghambatan dihitung sebagai total waktu yang dibutuhkan untuk menyelesaikan daftar konflik dikurangi total waktu yang dibutuhkan untuk menyelesaikan daftar tanpa konflik (Gilhooly & Fioratou, 2009; Miyake et al., 2000). Oleh karena itu, skor yang lebih tinggi mencerminkan biaya yang lebih besar untuk menangani konflik, dan karenanya kemampuan inhibisi yang lebih buruk.

### **8.3.3 Pergeseran**

Kemampuan berpindah diukur dengan tugas kategorisasi huruf-angka. Peserta melihat pasangan huruf-angka dan dalam satu daftar mengidentifikasi apakah huruf itu vokal atau konsonan, dalam daftar lain mengidentifikasi apakah jumlahnya ganjil atau genap, dan dalam daftar ketiga beralih di antara dua tugas. Setiap daftar terdiri dari 64 percobaan didahului oleh 5 percobaan praktek. Setiap stimulus ditampilkan hingga respons dengan jeda 500 ms di antara uji coba. Ukuran kemampuan pergeseran dihitung sebagai waktu respon rata-rata (RT) pada percobaan switching dikurangi rata-rata RT pada percobaan non-switching. Oleh karena itu, skor yang lebih tinggi mencerminkan kemampuan berpindah yang kurang efisien.

### **8.3.4 Penalaran Kondisional**

Tugas Inferensi Kondisional (Evans et al., 1995; Inglis & Simpson, 2009a) digunakan sebagai ukuran kemampuan penalaran untuk konsistensi dengan studi longitudinal di mana pengembangan bersama studi matematika ditemukan (lihat Bagian 4.3.2 untuk rincian tugas). Dalam studi ini, tugas itu diberikan oleh komputer sebagai lawan pena dan kertas. Item disajikan secara berurutan dan urutannya diacak di antara peserta. Setiap item tetap ditampilkan di layar hingga respons dengan jeda 500 ms di antara uji coba. Peserta menjawab dengan menekan tombol 'S' atau 'L' pada keyboard standar, masing-masing diberi label 'Tidak' dan 'Ya'. Empat indeks interpretasi, MCI, DCI, BCI dan CCI, dan dua indeks bias, NCI (ukuran NCB) dan API (ukuran APB), dihitung dari tugas seperti pada bab sebelumnya.

## **8.4 PROSEDUR**

Peserta mengambil bagian dalam laboratorium komputer dalam kelompok sekitar 30 orang. Mereka bekerja sendiri dan dalam keheningan. Tugas penalaran selalu diselesaikan terlebih dahulu sehingga sifat cepat dari tugas fungsi eksekutif tidak akan mengganggu kinerja pada tugas penalaran yang tidak dipercepat. Tugas fungsi eksekutif disajikan dalam urutan yang ditetapkan: memori kerja, penghambatan, lalu pergeseran. Sesi ini berlangsung sekitar 35 menit.

## **8.5 HASIL**

### **8.5.1 Pembersihan Data**

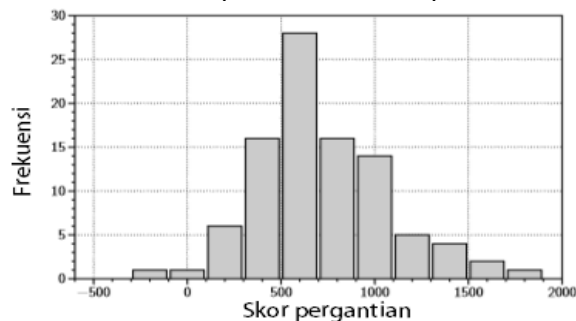
Skor satu peserta pada tugas penghambatan dihapus karena lebih dari tiga standar deviasi di atas rata-rata (mewakili kinerja yang sangat buruk). Semua data lainnya berada di antara rata-rata tugas dan  $\pm 3$  standar deviasi.

## **8.6 PERFORMA TUGAS**

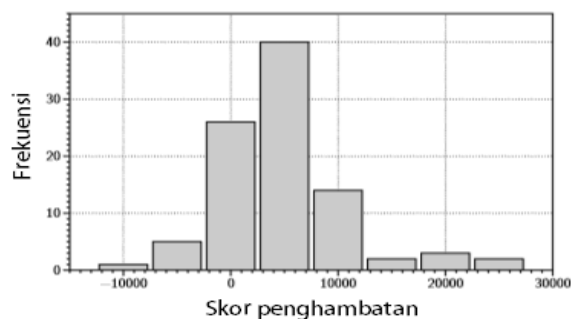
Kinerja di semua tugas seperti yang diharapkan. Perhatikan bahwa pada tugas memori kerja skor yang lebih tinggi mencerminkan kinerja yang lebih baik, sedangkan dengan tugas penghambatan dan pergeseran skor yang lebih tinggi mencerminkan biaya konflik/pergeseran yang lebih besar, dan karenanya keterampilan yang lebih buruk. Pergeseran skor,  $M = 712.09\text{ms}$ ,  $SD = 350.30$ , secara signifikan di atas nol,  $t(93) = 19.76$ ,  $p < .001$ , menunjukkan

bahwa rata-rata peserta lebih lambat pada uji coba switching dibandingkan uji coba non-switching, seperti yang diharapkan ( Gambar 9.2). Skor penghambatan,  $M = 4829.10$ ,  $SD = 6009.47$ , juga secara signifikan di atas nol,  $t(92) = 7.75$ ,  $p < .001$ , menunjukkan bahwa konflik warna kata/font memperlambat respons peserta, sekali lagi seperti yang diharapkan (Gambar 9.3 ). Terakhir, skor memori kerja (proporsi item yang dikategorikan dengan benar),  $M = 0,81$ ,  $SD = 0,08$ , juga secara signifikan di atas level 50%,  $t(93) = 39,02$ ,  $p < 0,001$  (Gambar 9.4). Kumpulan hasil ini menunjukkan bahwa peserta terlibat dengan tugas dan menunjukkan pola kinerja yang diharapkan dari literatur.

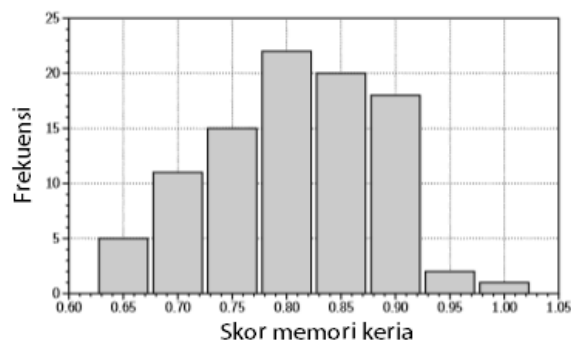
Pada tugas conditional inference, indeks interpretasi dengan mean tertinggi adalah BCI ( $M = 20.28$ ,  $SD = 5.99$ ), diikuti oleh CCI ( $M = 19.81$ ,  $SD = 3.13$ ), kemudian MCI ( $M = 19.47$ ,  $SD = 3,37$ ), dan terakhir DCI ( $M = 18,04$ ,  $SD = 5,61$ ). Hal ini tidak mengherankan mengingat temuan sebelumnya dalam penelitian ini bahwa mahasiswa non-matematika cenderung menunjukkan interpretasi yang lebih bikondisional daripada cacat interpretasi atau materi.



**Gambar 8.2:** Distribusi skor pada shift tugas.



**Gambar 8.3:** Distribusi skor pada tugas inhibisi (Stroop).



**Gambar 8.4:** Distribusi skor pada tugas memori kerja (2-belakang).

## 8.7 HUBUNGAN DI ANTARA FUNGSI EKSEKUTIF

Hubungan antara tiga ukuran fungsi eksekutif dianalisis dengan korelasi Pearson dan diringkas dalam Tabel 8.1. Hipopenelitian 1 menyatakan bahwa ketiga komponen fungsi eksekutif tidak akan saling berhubungan. Memori kerja tidak berkorelasi secara signifikan dengan kemampuan penghambatan,  $r(93) = .02, p = .827$ , juga tidak dengan kemampuan pergeseran,  $r(94) = .11, p = .294$ . Inhibisi dan skor pergeseran juga tidak berkorelasi secara signifikan,  $r(93) = .17, p = .104$ . Hal ini konsisten dengan temuan sebelumnya bahwa tiga komponen fungsi eksekutif adalah keterampilan yang dapat dipisahkan (Ardila et al., 2000; Arffa, 2007; Friedman et al., 2006; Handley et al., 2004; Miyake et al., 2000).<sup>1</sup>

**Tabel 8.1:** Korelasi antar Fungsi Eksekutif. Semua  $p > .1$ .

	Halangan	Pergantian
Memori pekerjaan	.02	-.11
Halangan		.17

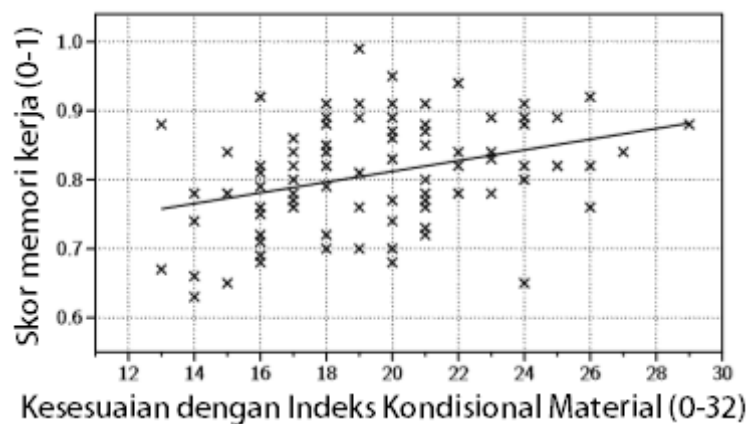
**Tabel 8.2:** Korelasi antara Fungsi Eksekutif dan indeks interpretasi. <sup>+</sup> $p < .1$ , \* $p < .05$ , \*\* $p < .01$ .

	Halangan	Pergantian	Memori Pengerjaan
MCI	-.13	-.12	.34**
DCI	-.18 <sup>+</sup>	-.18 <sup>+</sup>	.12
CCI	-.13	.01	-.12
BCI	.21*	.22*	-.08

## 8.8 FUNGSI EKSEKUTIF DAN PENALARAN KONDISIONAL.

### 8.8.1 Indeks Interpretasi

Hubungan antara tiga ukuran fungsi eksekutif dan empat indeks interpretasi diselidiki dengan korelasi Pearson dan diringkas dalam Tabel 8.2. Hipopenelitian 2 menyatakan bahwa skor memori kerja akan berkorelasi positif dengan MCI, dan ini didukung oleh data,  $r(94) = .34, p = .001$ . Mereka yang memiliki memori kerja yang lebih baik memiliki skor MCI yang lebih tinggi (Gambar 8.5).



**Gambar 8.5:** Korelasi antara MCI dan skor memori kerja.

Hipopenelitian 3 menyatakan bahwa skor inhibisi tidak akan berhubungan dengan indeks interpretasi karena sifatnya yang abstrak. Berlawanan dengan ini, skor penghambatan secara signifikan berkorelasi positif dengan BCI,  $r(93) = .21, p = .041$ , dan sedikit berkorelasi

negatif dengan DCI,  $r(93) = .18, p = .083$ , menunjukkan bahwa penghambatan yang lebih baik (diwakili oleh skor yang lebih rendah) dikaitkan dengan BCI yang lebih rendah dan DCI yang sedikit lebih tinggi. Kedua koefisien korelasi ini sedikit berbeda,  $t(91) = 1,93, p = 0,056$ . Ini bisa jadi karena pada non-matematikawan interpretasi bikondisional berasal dari tingkat kognisi heuristik, dan ketika dihambat digantikan dengan interpretasi kondisional yang rusak. Dalam penelitian sebelumnya tidak ada indikasi hubungan antara penghambatan dan perilaku penalaran kondisional abstrak, tetapi ini mungkin karena penelitian sebelumnya hanya melihat konsistensi dengan interpretasi materi dan tidak mempertimbangkan BCI atau DCI. Mungkin saja mengambil interpretasi material tergantung pada faktor lain (misalnya memori kerja) sedangkan BCI dan DCI terkait dengan penghambatan dengan cara yang disarankan di atas.

Tidak ada prediksi pasti tentang hubungan antara pergeseran dan penalaran kondisional, tetapi pergeseran ditemukan berkorelasi positif secara signifikan dengan BCI,  $r(94) = .22, p = .036$ , dan sedikit berkorelasi negatif dengan DCI,  $r(94) = .18, p = .077$ . Ini menunjukkan bahwa kemampuan pergeseran yang lebih baik (diwakili oleh skor yang lebih rendah) dikaitkan dengan interpretasi kondisional yang lebih sedikit, dan interpretasi yang sedikit lebih cacat, mencerminkan hubungan yang ditemukan untuk penghambatan.

### **8.8.2 Bias dalam Penalaran Kondisional**

Terakhir, hubungan antara tiga fungsi eksekutif peserta dan NCI dan API mereka diperiksa, sekali lagi dengan korelasi Pearson. Hipopenelitian 4 menyatakan bahwa NCB akan terkait dengan inhibisi atau skor memori kerja, dan Hipopenelitian 5 menyatakan bahwa APB akan terkait dengan skor inhibisi.

NCI secara signifikan berkorelasi negatif dengan skor pergeseran,  $r(93) = .26, p = 0,013$ , menunjukkan, agak berlawanan dengan intuisi, bahwa mereka dengan kemampuan pergeseran yang lebih baik memiliki NCI yang lebih tinggi. Kemungkinan yang menarik adalah hubungan kuadrat antara DCI dan NCI yang ditemukan di Bab 4 didorong di beberapa bagian oleh keterampilan pergeseran, yaitu kemampuan pergeseran yang lebih baik dikaitkan dengan NCI yang lebih tinggi dan interpretasi yang lebih cacat dari kondisi (secara signifikan signifikan di sini) yang berarti bahwa NCI cenderung lebih tinggi pada orang dengan interpretasi kondisional yang lebih cacat, setidaknya sampai titik tertentu.

NCI tidak berkorelasi dengan skor memori kerja,  $r(93) = .08, p = .472$ , atau dengan skor penghambatan,  $r(92) = .07, p = .493$ , bertentangan dengan Hipopenelitian 4. API secara signifikan negatif berkorelasi dengan kemampuan menggeser,  $r(94) = .21, p = .043$ , di mana mereka yang memiliki kemampuan menggeser lebih baik menunjukkan API yang lebih tinggi, mirip dengan NCI. API sedikit berkorelasi negatif dengan penghambatan,  $r(92) = .19, p = .066$ , menunjukkan bahwa mereka yang memiliki kemampuan penghambatan yang lebih baik cenderung menunjukkan API yang lebih besar, bertentangan dengan prediksi. Terakhir, API tidak berkorelasi dengan skor memori kerja,  $r(93) = .07, p = .485$ .

## **8.9 DISKUSI**

Tujuan dari bab ini adalah untuk menyelidiki potensi tingkat kognisi algoritmik menjadi sumber perubahan dalam perilaku penalaran kondisional. Disarankan bahwa level algoritmik mungkin merupakan mekanisme melalui mana TFD dapat beroperasi, tetapi untuk kasus ini, perlu ada hubungan antara ukuran level algoritmik dan ukuran perilaku penalaran. Studi yang disajikan dalam Bab 5 menyarankan bahwa kecerdasan, pada tingkat algoritmik, bukanlah

mekanisme untuk perubahan yang ditemukan dalam perilaku penalaran mahasiswa matematika. Di sini, tiga fungsi eksekutif dari memori kerja, penghambatan dan pergeseran digunakan sebagai ukuran lebih lanjut dari tingkat algoritmik. Tugas Inferensi Kondisional yang digunakan dalam Bab 4 dan 5, yang ditemukan mahasiswa matematika berubah secara signifikan selama tingkat AS, digunakan sebagai ukuran kemampuan penalaran. Kinerja pada tugas fungsi eksekutif seperti yang diharapkan: distribusi skor sejalan dengan studi sebelumnya dan ketiga keterampilan tersebut ternyata dapat dipisahkan dengan jelas (Hipopenelitian 1).

Hipopenelitian 2 meramalkan bahwa memori kerja akan berkorelasi positif dengan MCI, dan inilah masalahnya. Memori kerja yang lebih baik dikaitkan dengan interpretasi yang lebih material dari kondisional, mengkonfirmasi temuan penelitian sebelumnya (misalnya De Neys et al., 2005; Verschueren et al., 2005). Interpretasi lain tampaknya tidak terkait dengan memori kerja dan MCI tidak berkorelasi dengan inhibisi atau keterampilan bergeser. Ini bisa mencerminkan bahwa, daripada mengandalkan penghambatan tanggapan intuitif mendukung interpretasi alternatif yang tersedia, interpretasi material bergantung pada memiliki kapasitas memori kerja yang diperlukan untuk menerapkan atau menghitung alternatif.

Penghambatan dan pergeseran ditemukan berkorelasi positif dengan BCI dan sedikit berkorelasi negatif dengan DCI. Ini bisa menunjukkan bahwa BCI adalah interpretasi intuitif dari pernyataan kondisional, dan bahwa mereka yang memiliki kemampuan penghambatan dan pergeseran yang lebih baik lebih mampu menghambat respons ini demi interpretasi yang rusak. Hal ini berimplikasi pada TFD: ada kemungkinan bahwa ketika mahasiswa matematika menjadi kurang bikondisional dan lebih cacat dalam pola penalaran mereka (seperti yang ditemukan pada mahasiswa tingkat AS di Bab 4), itu karena keterampilan penghambatan dan pergeseran mereka telah meningkat. Hal ini tidak bertentangan dengan hipopenelitian yang diajukan dalam Bab 4 yang menyarankan bahwa paparan ke depan ('jika kemudian') inferensi mendorong mahasiswa untuk mempertimbangkan kasus tidak-p tidak relevan. Mungkin paparan inilah yang melatih mahasiswa matematika untuk menghambat interpretasi bikondisional.

Hasil yang terkait dengan skor NCI dan API sangat mengejutkan. Untuk NCI dihipopenelitiankan bahwa baik penghambatan (jika NCB adalah bias Tipe 1) atau skor memori kerja (jika NCB adalah bias Tipe 2) akan berkorelasi, namun keduanya tidak. Untuk API dihipopenelitiankan bahwa skor penghambatan akan berkorelasi, dan ternyata tidak. Namun, mungkin jenis penghambatan yang diukur di sini bukan jenis penghambatan yang sama yang diperlukan untuk menghindari NCB dan APB. Dalam tugas fungsi eksekutif peserta diberitahu bagaimana melakukan secara optimal dan ukuran mencerminkan seberapa baik mereka dapat melakukannya (Stanovich, 2009a). Dalam kasus tugas penghambatan yang digunakan di sini, peserta diberitahu untuk menghambat gangguan warna font pada tanggapan mereka dan ukurannya adalah biaya RT untuk melakukannya. Penghambatan spontan kesalahan Tipe 1 oleh intervensi pemrosesan Tipe 2 saat menyelesaikan tugas penalaran mungkin memiliki sifat yang sama sekali berbeda. Hal ini mungkin tergantung pada disposisi bernalar dari orang yang bernalar pada tingkat kognisi reflektif daripada efisiensi penerapan inhibisi pada tingkat algoritmik (Stanovich, 2009a). Cara yang lebih efektif untuk menguji hubungan antara penghambatan dan NCB dan APB adalah dengan menggunakan ukuran penghambatan tingkat reflektif seperti Uji Refleksi Kognitif (Toplak et al., 2011).



Analisis seperti itu dimungkinkan dalam data yang disajikan dalam Bab 6 di mana satu set peserta sarjana menyelesaikan CRT dan Tugas Inferensi Kondisional. Analisis ulang data ini menunjukkan bahwa skor intuitif terbalik pada CRT secara signifikan berkorelasi negatif dengan APB,  $r(55) = .30, p = .028$ , menunjukkan bahwa mereka yang memiliki penghambatan respons intuitif yang lebih baik terhadap CRT memiliki skor APB yang lebih rendah. . Sebaliknya, tidak ada korelasi yang signifikan antara skor intuitif terbalik dengan skor CRT dan NCB,  $r(55) = .18, p = .187$ . Analisis tambahan ini menunjukkan bahwa APB memang bias heuristik, sementara tidak ada bukti bahwa hal ini terjadi pada NCB.

Penting untuk dicatat bahwa hasil bab ini tidak mendukung gagasan bahwa fungsi eksekutif bertanggung jawab atas perubahan pada perilaku penalaran mahasiswa matematika, hanya saja hal itu bisa terjadi. Studi longitudinal yang dilaporkan dalam Bab 4 menyarankan bahwa disposisi bernalar (diukur dengan skor CRT) bukan merupakan faktor dalam perubahan perilaku penalaran, tetapi itu adalah prediktor kemampuan penalaran awal. Mungkin kasus pergeseran dan penghambatan memainkan peran yang sama, mempengaruhi kemampuan penalaran dasar seseorang tetapi tidak memainkan peran dalam pengembangan di samping studi matematika. Untuk menyelidiki ini lebih lanjut, studi longitudinal lain akan diperlukan di mana fungsi eksekutif mahasiswa matematika dan non-matematika diukur pada awal dan akhir periode studi. Namun demikian, hasil dari bab ini memberikan beberapa bukti bahwa fungsi eksekutif dapat menjadi tempat yang menarik untuk mencari mekanisme perubahan penalaran.

### **8.10 RINGKASAN DAN TEMUAN BARU**

1. Penghambatan dan pergeseran berkorelasi positif dengan kesesuaian dengan interpretasi bikondisional dan sedikit berkorelasi negatif dengan kesesuaian dengan interpretasi kondisional yang cacat dari pernyataan kondisional.
2. Bias premis afirmatif tidak terkait dengan ukuran inhibisi fungsi eksekutif, tetapi terkait dengan ukuran inhibisi tingkat reflektif, CRT, dalam analisis ulang kumpulan data sebelumnya.

## **BAB 9**

### **SIMPULAN**

#### **9.1 PENDAHULUAN**

Selama ribuan tahun telah diasumsikan bahwa belajar matematika meningkatkan keterampilan penalaran umum. Gagasan ini dikenal dengan Teori Disiplin Formal (TFD). Penelitian sebelumnya menunjukkan bahwa mungkin ada hubungan antara matematika tingkat lanjut dan keterampilan penalaran kondisional (Lehman & Nisbett, 1990; Inglis & Simpson, 2008), meskipun masih belum jelas apakah hubungan tersebut bersifat perkembangan.

Penelitian dalam buku ini bertujuan untuk melakukan dua hal: menetapkan apakah belajar matematika pada tingkat lanjutan dikaitkan dengan perubahan keterampilan penalaran, dan jika demikian, untuk menyelidiki beberapa mekanisme potensial untuk perubahan yang ditemukan. Temuan studi saya menunjukkan bahwa belajar matematika dikaitkan dengan kurang bikondisional dan lebih banyak materi, dan khususnya, lebih cacat, interpretasi pernyataan kondisional. Ini tampaknya terbatas pada penalaran kondisional abstrak 'jika kemudian'. Sumber perubahan dibahas lebih lanjut di bawah ini, setelah ringkasan dari setiap temuan studi.

#### **9.2 GAMBARAN UMUM INTERPRETASI DAN PENEMUAN**

Bab 4 mendokumentasikan studi longitudinal satu tahun dengan mahasiswa yang mengambil matematika dan bahasa tingkat AS. Di Indonesia, tingkat AS adalah tahap pertama dari studi pasca-wajib, dan mahasiswa dapat memilih mata pelajaran mana yang ingin mereka pelajari (biasanya empat). Para mahasiswa yang telah memilih untuk belajar matematika dan mereka yang telah memilih bahasa tidak berbeda pada tugas Inferensi Kondisional (memvalidasi pemotongan dari pernyataan bentuk 'jika p maka q') pada awal studi mereka. Namun, pada akhir tahun studi mereka, mahasiswa matematika telah menjadi lebih sedikit bikondisional dan secara signifikan lebih banyak materi dan cacat dalam interpretasi mereka terhadap pernyataan kondisional, sementara mahasiswa bahasa tidak berubah.

Pemeriksaan ukuran efek mengungkapkan bahwa perubahan indeks kondisional cacat ( $d = 0,880$ ) terutama lebih besar daripada salah satu dari perubahan lainnya. Karena kecerdasan umum dan disposisi bernalar dikendalikan, tampaknya perubahan ini didorong oleh studi matematis peserta. Tampaknya belajar matematika mengajarkan mahasiswa untuk mengasumsikan p, dan karena itu menganggap kasus bukan-p tidak relevan. Sebagaimana dibahas dalam Bab 1.3 tentang rasionalitas, interpretasi kondisional material adalah standar normatif (dianggap benar oleh ahli logika), dan interpretasi yang cacat, meskipun bukan model normatif, dapat dianggap sebagai 'lebih baik' daripada interpretasi bikondisional. Dalam pengertian ini dapat dikatakan bahwa mahasiswa matematika meningkat dalam penalaran mereka dengan pernyataan kondisional. Mereka tidak, bagaimanapun, mengubah penalaran mereka dengan silogisme tematik, yang selanjutnya mendukung gagasan bahwa perubahan itu dalam interpretasi mahasiswa tentang 'jika'.

Penjelasan yang diusulkan untuk temuan ini adalah matematika secara teratur menghadapkan mahasiswa pada pernyataan kondisional implisit (silabus tingkat AS tidak

berisi referensi eksplisit untuk logika kondisional) di mana mereka diharapkan untuk mengasumsikan  $p$  dan alasan tentang  $q$ . Houston (2009) berpendapat bahwa sebagian besar pernyataan matematis berbentuk 'jika pernyataan A benar, maka pernyataan B benar', bahkan jika pernyataan tersebut sangat disamarkan (hal. 63). Dia juga berpendapat bahwa dalam matematika A dianggap benar, bahkan ketika itu jelas tidak benar (misalnya, dalam kasus bukti kontradiksi), dan kebenaran atau kesalahan B kemudian disimpulkan. Untuk alasan ini, Hoyles dan Kuchemann (2002) berpendapat bahwa interpretasi yang cacat sebenarnya lebih tepat untuk kelas matematika daripada interpretasi materi. Senada dengan penalaran tersebut, Inglis dan Simpson (2009b) menemukan bahwa sekelompok mahasiswa sarjana matematika yang sukses cenderung memiliki lebih banyak kekurangan daripada materi interpretasi kondisional.

Studi level AS dimodifikasi dan diulangi dengan mahasiswa sarjana tahun pertama di Bab 5, dan pola perubahan direplikasi. Interpretasi mahasiswa matematika terhadap kondisi menjadi lebih banyak materi ( $d = .55$ ), lebih banyak cacat ( $d = .66$ ), dan lebih sedikit kondisional ( $d = .46$ ) dari waktu ke waktu. Namun, mahasiswa psikologi menunjukkan pola perubahan yang sama dan ada juga masalah kekuatan statistik yang rendah. Akibatnya, pola perubahan dari mahasiswa tingkat AS hanya dapat dideteksi pada mahasiswa sarjana dengan uji-t yang membandingkan setiap indeks interpretasi sepanjang waktu.

Alih-alih tugas silogisme tematik, para peserta dalam Bab 6 menyelesaikan tugas Inferensi Kondisional tematik. Hal ini memungkinkan luasnya perubahan dalam interpretasi yang cacat untuk diselidiki. Menariknya, mahasiswa matematika tidak menunjukkan perubahan dalam indeks interpretasi dari waktu ke waktu. Meskipun kekuatan statistik kembali menjadi masalah, bahkan tidak ada tren untuk perubahan. Hal ini menunjukkan bahwa belajar matematika hanya dapat mengubah interpretasi mahasiswa terhadap pernyataan kondisional abstrak.

Dalam Bab 6, mahasiswa matematika sarjana tahun ketiga menyelesaikan tugas Inferensi Kondisional dengan ungkapan kondisional sebagai 'jika  $p$  maka  $q$ ' atau sebagai ' $p$  hanya jika  $q$ '. Kedua bentuk kondisional ini secara logis setara, tetapi penelitian sebelumnya telah menunjukkan mahasiswa non-matematika memperlakukan mereka secara berbeda (Evans, 1977). Bab 4 dan 5 menyarankan bahwa efek belajar matematika pada keterampilan penalaran mungkin terbatas pada pernyataan kondisional abstrak. Dihipotesiskan bahwa ini bisa jadi karena mahasiswa berulang kali dihadapkan pada pernyataan (implisit) 'jika  $p$  maka  $q$ ' di mana mereka mengasumsikan bahwa  $p$  benar dan beralasan tentang  $q$ .

Evans (1977) berpendapat bahwa pernyataan 'jika kemudian' dan 'hanya jika' diperlakukan berbeda karena kebutuhan dan kecukupan ditekankan secara berbeda di masing-masing pernyataan. Sementara pernyataan 'jika  $p$  maka  $q$ ' menekankan kecukupan  $p$ , ' $p$  hanya jika  $q$ ' menekankan perlunya  $q$ . Jika mahasiswa matematika belajar untuk berasumsi bahwa 'jika  $p$  maka  $q$ ' berarti bahwa  $p$  benar dan mereka perlu bernalar tentang  $q$ , yang bertentangan dengan interpretasi umum yang cacat dari pernyataan kondisional, maka perlunya  $q$  ditekankan dalam 'hanya jika' pernyataan dapat mengganggu penalaran mereka dan menyebabkan mereka untuk merespon kurang cacat. Ini adalah efek yang ditemukan.

Mahasiswa yang menyelesaikan versi 'hanya jika' dari tugas merespons lebih bikondisional dan kurang cacat dan material dibandingkan rekan-rekan mereka dalam kondisi 'jika kemudian'. Tingkat dukungan mahasiswa matematika dalam kondisi 'hanya jika' tampaknya mirip dengan tingkat dukungan mahasiswa non-matematika dalam penelitian

Evans (1977). Ini menunjukkan bahwa cacat interpretasi mahasiswa matematika dari kondisi terbatas pada pernyataan 'jika kemudian', dan mungkin memang berasal dari paparan daripada dari perubahan umum dalam pemahaman mereka tentang kondisional.

Potensi tingkat kognisi heuristik menjadi mekanisme perubahan kemampuan penalaran mahasiswa matematika diselidiki pada Bab 7. Mahasiswa sarjana matematika dan non-matematika menyelesaikan tugas Inferensi Kondisional abstrak standar dua kali: sekali dengan batas waktu yang sangat singkat untuk setiap item dan sekali dengan waktu sebanyak yang mereka suka. Batas waktu terbukti cukup lama bagi tingkat heuristik untuk menghasilkan respons tetapi cukup pendek untuk mencegah penalaran Tipe 2 mengganggu. Di bawah batasan ini, mahasiswa matematika merespons secara signifikan kurang sejalan dengan interpretasi yang cacat dari kondisional daripada di bawah kondisi waktu luang, tetapi mereka masih sedikit lebih cacat daripada mahasiswa non-matematika. Hal ini menunjukkan bahwa interpretasi yang cacat memang berasal, sebagian, dari tingkat heuristik, tetapi penalaran Tipe 2 memainkan peran penting dalam mendorongnya.

Temuan ini sesuai dengan hipotesis yang diajukan di atas, bahwa mahasiswa matematika dihadapkan pada pernyataan implisit 'jika maka' yang berarti mereka belajar untuk mengasumsikan  $p$  dan sebagai konsekuensinya memiliki pemahaman yang salah tentang kondisi. Paparan berulang terhadap stimulus implisit dari waktu ke waktu dapat membawa perubahan pada tingkat heuristik, dengan cara yang sama bahwa pemahaman kata dan makna secara umum menjadi otomatis ketika kita belajar membaca (LaBerge & Samuels, 1974). Sebuah kebiasaan bawah sadar untuk mempertimbangkan tidak- $p$  kasus tidak relevan dapat diperkuat oleh Tipe 2 bernalar melalui proses pembenaran (Evans, 2006, 2011), di mana output dari tingkat heuristik dibenarkan daripada diteliti. Hal ini pada gilirannya dapat mendorong kecenderungan cacat dalam masalah berikutnya ("Saya menjawab seperti ini sebelumnya jadi saya mungkin bernalar sama tentang pertanyaan ini"), yang mengarah ke Indeks Kondisi Cacat yang lebih tinggi dalam kondisi lambat.

Hal ini membuka kemungkinan bahwa meningkatkan refleksi sadar atau menambahkan 'aturan' (seperti definisi materi kondisional) pada akhirnya menyebabkan mahasiswa matematika bergerak lebih ke arah respon kondisional materi di kemudian hari dalam pendidikan matematika mereka. Penjelasan alternatif adalah matematikawan materi sebenarnya memiliki interpretasi yang cacat tetapi mampu menggunakan bukti kontradiksi untuk mendukung inferensi MT, membuat tanggapan mereka sesuai dengan kondisi materi.

Akhirnya, Bab 8 menyelidiki peran fungsi eksekutif (pada tingkat kognisi algoritmik) dalam penalaran kondisional. Meskipun peserta adalah mahasiswa non-matematika, ini memungkinkan indikasi apakah ada potensi fungsi eksekutif (ditunjukkan oleh hubungan dengan perilaku penalaran kondisional) menjadi mekanisme perubahan. Pertama, ditunjukkan bahwa memori kerja terkait dengan indeks kondisional material, yang mereplikasi temuan sebelumnya. Temuan baru adalah penghambatan dan pergeseran yang lebih baik terkait dengan respons bikondisional yang lebih sedikit dan respons yang sedikit lebih rusak. Disarankan bahwa ini bisa jadi karena interpretasi bikondisional adalah interpretasi intuitif, dan ketika ini dapat dihambat, itu cenderung digantikan oleh interpretasi yang rusak.

Hal ini tampaknya bertentangan dengan temuan dari Bab 7, dengan menyarankan bahwa interpretasi yang rusak diimplementasikan pada tingkat penalaran Tipe 2 daripada pada tingkat heuristik. Namun, proposal yang diuraikan di atas adalah belajar matematika mendorong interpretasi yang rusak ke tingkat heuristik, dan peserta dalam studi fungsi

eksekutif berasal dari program gelar non-matematika. Bisa jadi dalam populasi umum, interpretasi bikondisional adalah interpretasi yang intuitif, dan bahwa peserta IQ tinggi mampu merespons dengan lebih tidak sempurna (Evans et al., 2007) karena mereka lebih berhasil dalam menghambat respons heuristik dan menerapkan Tipe 2 respon tingkat. Dalam matematikawan, interpretasi bikondisional mungkin secara bertahap digantikan oleh interpretasi yang cacat sebagai standar intuitif, dan ini hanya diperkuat pada tingkat penalaran Tipe 2.

Temuan studi fungsi eksekutif menunjukkan bahwa perubahan matematika awalnya mungkin terjadi melalui penghambatan yang lebih besar dari interpretasi bikondisional. Mungkin paparan pernyataan 'jika kemudian' mendorong mahasiswa untuk menghambat interpretasi bikondisional karena terbukti tidak sesuai melalui contoh yang mereka hadapi. Pengalaman berulang dari asumsi  $p$  dan penalaran tentang  $q$  juga mulai menggantikan interpretasi bikondisional yang dipegang dalam pikiran heuristik. Akhirnya, interpretasi yang rusak menjadi intuitif dan selanjutnya diperkuat dengan penalaran Tipe 2 melalui proses membenaran. Pada titik tertentu, instruksi eksplisit dalam logika kondisional dapat memberikan mahasiswa dengan pengetahuan yang diperlukan untuk interpretasi materi, dan penghambatan mungkin lagi berperan dalam menghambat interpretasi yang rusak mendukung interpretasi materi yang baru dipelajari.

### 9.3 PENELITIAN LANJUTAN

Penelitian yang disajikan di sini telah membuka beberapa pertanyaan yang memerlukan penelitian lebih lanjut. Pertama, tidak ada bukti bahwa mempelajari matematika memiliki dampak pada penalaran di luar penalaran kondisional 'jika kemudian' abstrak. Namun demikian, akan bermanfaat untuk menyelidiki lebih lanjut jenis keterampilan penalaran deduktif, seperti penalaran disjungtif, dan keterampilan penalaran informal, seperti penerapan hukum bilangan besar dan keterampilan lain yang dijelaskan dalam Bab 4. Ada kemungkinan bahwa bidang matematika yang berbeda berdampak pada berbagai jenis keterampilan penalaran, dan beberapa keterampilan yang terpengaruh tidak diukur dengan tugas kondisional dan silogisme yang digunakan di sini.

Dalam nada yang sama, juga akan bermanfaat untuk mengisolasi dampak dari berbagai bidang matematika pada keterampilan penalaran kondisional abstrak. Mungkin saja, misalnya, pembelajaran geometri tidak terkait dengan perubahan perilaku penalaran kondisional sementara pembelajaran kalkulus. Upaya pertama pada tugas yang sulit ini bisa untuk menguji pengembangan keterampilan penalaran kondisional di berbagai kurikulum, misalnya, di Siprus di mana ada penekanan besar pada matematika berbasis geometri.

Sebuah jalan penting dari penyelidikan untuk pekerjaan masa depan adalah untuk melihat apakah temuan dari studi tingkat AS akan mereplikasi dalam sistem pendidikan di mana wajib bagi mahasiswa untuk belajar beberapa bentuk matematika sampai usia 18 tahun. Para mahasiswa di tingkat AS (dan sarjana) studi telah memilih untuk tinggal di pendidikan dan telah memilih mata pelajaran yang akan dipelajari. Bisa jadi disposisi ini penting untuk efek yang ditemukan. Ini hasil dari desain eksperimen semu, yang tidak dapat dihindari dalam kasus ini.

Namun, mereplikasi studi dalam kohort wajib-matematika akan membantu mengurangi kebingungan keinginan untuk belajar matematika. Berdasarkan temuan dari penelitian ini disarankan bahwa mungkin ada tren perkembangan dalam hubungan antara

belajar matematika dan perubahan perilaku penalaran kondisional, di mana mahasiswa beralih dari memiliki pandangan kondisional ke pandangan cacat, dan akhirnya ke pandangan materi. Tren ini, bersama dengan mekanisme yang diusulkan (perubahan pada tingkat kognisi heuristik yang diperkuat pada tingkat kognisi algoritmik), harus diuji. Hal ini dapat dilakukan baik dengan studi longitudinal mengikuti mahasiswa dari tingkat AS hingga akhir tingkat gelar, dan jika memungkinkan, ke studi pascasarjana. Alternatifnya, dan jauh lebih ekonomis, tren dapat diselidiki dengan studi cross-sectional yang membandingkan kelompok ahli matematika dan non-matematika dari setiap tahap antara tingkat AS dan tingkat staf akademik. Ini akan memungkinkan kita untuk melihat apakah, dan pada tahap mana, matematikawan berubah dari memiliki interpretasi yang cacat dari kondisi ke interpretasi material dari kondisi.

#### **9.4 MENINJAU KEMBALI TEORI DISPLIN FORMAL**

Penelitian ini didorong oleh TFD, yang mengklaim, tanpa bukti, bahwa belajar matematika meningkatkan keterampilan penalaran umum seseorang. Bukti yang ditemukan di sini sangat terbatas dalam mendukung TFD. Sementara mahasiswa matematika melakukan 'meningkatkan' dalam penalaran kondisional abstrak berdasarkan pernyataan 'jika kemudian', mereka tidak menjadi lebih normatif secara langsung, dan mereka tidak ditemukan meningkatkan pada penalaran tematik 'jika kemudian', penalaran kondisional abstrak dari bentuk 'hanya jika', atau pada silogisme tematik. Berdasarkan bukti ini, TFD tampaknya telah dlebih-lebihkan di masa lalu. Kutipan seperti "Melalui matematika kami juga ingin mengajarkan penalaran logis - tidak ada alat yang lebih baik untuk itu telah ditemukan sejauh ini" dari Amitsur (Sfard, 1998, hal. 453) dan "Studi matematika tidak dapat digantikan oleh aktivitas lain yang akan melatih dan mengembangkan kemampuan logika murni manusia ke tingkat rasionalitas yang sama" dari Oakley (1949, hlm. 19) tidak didukung oleh penelitian yang disajikan di sini: istilah 'penalaran logis' dan 'rasionalitas' pasti dimaksudkan untuk merujuk pada perilaku melampaui penalaran 'jika maka' abstrak. Hal ini juga berimplikasi pada kebijakan pendidikan matematika. TFD telah digunakan sebagai argumen untuk matematika untuk diprioritaskan dalam Kurikulum Nasional Indonesia dalam beberapa cara, misalnya oleh Smith (2004), yang mengatakan bahwa belajar matematika "mendisiplinkan pikiran, mengembangkan penalaran logis dan kritis, dan mengembangkan analitis dan keterampilan pemecahan masalah ke tingkat yang tinggi" (hal. 11). Klaim lebih lanjut dari sifat ini mungkin sebaiknya ditahan sampai mereka dapat didukung oleh bukti.

## DAFTAR LAMPIRAN

### 1. LAMPIRAN A

#### 1.1 Matriks Progresif Tingkat Lanjut

Subset Matriks Progresif Tingkat Lanjut Raven yang digunakan di seluruh penelitian ini terdiri dari item 12 dan 14-30 dari Set II. Instruksi yang digunakan disajikan di bawah ini. Peserta diberikan waktu 15 menit untuk menyelesaikan tugas pada Bab 5 dan 6 dan 10 menit pada Bab 8.

#### 1.2 Penyelesaian Pola

Di bagian ini Anda akan melihat kisi-kisi dan masing-masing akan memiliki pola dengan bagian yang hilang. Tugas Anda adalah memutuskan kepingan bernomor mana yang menyelesaikan kisi dengan benar. Dalam setiap soal, lingkari bagian yang menurut Anda benar. Ada 18 item dan Anda akan memiliki 15 atau 10 menit untuk melakukan sebanyak yang Anda bisa. Pertama, perhatikan contoh soal di bawah ini. Potongan berlabel 8 benar karena cocok dengan pola baik di bawah maupun di seberang kisi.

### 2. LAMPIRAN B

#### 2.1 Tugas Inferensi Kondisional

Pada halaman berikut adalah tugas inferensi kondisional penuh, termasuk instruksi, seperti yang disajikan kepada peserta (tetapi tanpa pengacakan item).

#### 2.2 Logika

***Silakan baca petunjuk berikut dengan seksama.***

Bagian ini berkaitan dengan kemampuan orang untuk bernalar secara logis dengan kalimat dalam berbagai bentuk. Disini akan disajikan total 32 masalah untuk Anda. Dalam setiap kasus Anda diberikan dua pernyataan bersama dengan kesimpulan yang mungkin atau tidak mungkin mengikuti pernyataan.

Tugas Anda dalam setiap kasus adalah memutuskan apakah kesimpulan harus mengikuti atau tidak dari pernyataan. Sebuah kesimpulan diperlukan jika itu harus benar, mengingat bahwa pernyataan itu benar.

Setiap soal menyangkut pasangan huruf-angka imajiner dan berisi pernyataan atau aturan awal yang menentukan huruf mana yang boleh dipasangkan dengan angka mana. Contoh aturan yang bentuknya mirip dengan yang digunakan adalah:

- Jika hurufnya B maka angkanya bukan 7.

Dalam setiap kasus Anda harus mengasumsikan bahwa aturan tersebut berlaku dan kemudian menggabungkannya dengan informasi yang diberikan dalam pernyataan kedua. Ini akan menyangkut huruf atau nomor pasangan imajiner, misalnya:

- Hurufnya itu adalah Y
- Angkanya bukan 4.

Jika informasinya menyangkut huruf, kesimpulannya akan menyangkut nomornya dan sebaliknya.

Masalah lengkap terlihat seperti:

Jika hurufnya X maka angkanya adalah 1.  
Hurufnya adalah X.

Kesimpulan: Angkanya adalah 1.

- YA
- TIDAK

Jika menurut Anda kesimpulannya harus mengikuti maka silakan centang kotak YA, jika tidak centang kotak TIDAK. Silakan selesaikan masalah secara berurutan dan pastikan Anda tidak melewatkannya. Jangan kembali ke masalah setelah Anda selesai dan pindah ke yang lain.

Jika menurut Anda kesimpulannya harus mengikuti, silakan centang YA, jika tidak centang TIDAK. Jangan kembali ke masalah setelah Anda selesai dan pindah ke yang lain. Jawab semua pertanyaan.

1. Jika hurufnya A maka angkanya 3.  
Hurufnya A.  
Kesimpulan: Angkanya 3.
  - a. YA
  - b. TIDAK
2. Jika hurufnya T maka angkanya 5.  
Hurufnya bukan T.  
Kesimpulan: Angkanya bukan 5.
  - a. YA
  - b. TIDAK
3. Jika hurufnya F maka angkanya 8.  
Angkanya 8.  
Kesimpulan: Hurufnya F.
  - a. YA
  - b. TIDAK
4. Jika hurufnya D maka angkanya 4.  
Angkanya bukan 4.  
Kesimpulan: Hurufnya bukan D.
  - a. YA
  - b. TIDAK
5. Jika hurufnya G maka angkanya bukan 6.  
Hurufnya G.  
Kesimpulan: Angkanya bukan 6.
  - a. YA
  - b. TIDAK
6. Jika hurufnya R maka angkanya bukan 1.  
Hurufnya bukan R.  
Kesimpulan: Angkanya 1.
  - a. YA
  - b. TIDAK
7. Jika hurufnya K maka angkanya bukan 3.  
Angkanya bukan 3.  
Kesimpulan: Hurufnya K.
  - a. YA
  - b. TIDAK



8. Jika hurufnya U maka angkanya bukan 9.  
Angkanya 9.  
Kesimpulan: Hurufnya bukan U.
  - a. YA
  - b. TIDAK
9. Jika hurufnya bukan B maka angkanya 5.  
Hurufnya bukan B.  
Kesimpulan: Angkanya 5.
  - a. YA
  - b. TIDAK
10. Jika hurufnya bukan S maka angkanya 6.  
Hurufnya S.  
Kesimpulan: Angkanya bukan 6.
  - a. YA
  - b. TIDAK
11. Jika hurufnya bukan V maka angkanya 8.  
Angkanya 8.  
Kesimpulan: Hurufnya bukan V.
  - a. YA
  - b. TIDAK
12. Jika hurufnya bukan H maka angkanya 1.  
Angkanya bukan 1.  
Kesimpulan: Hurufnya H.
  - a. YA
  - b. TIDAK
13. Jika hurufnya bukan F maka angkanya bukan 3.  
Hurufnya bukan F.  
Kesimpulan: Angkanya bukan 3.
  - a. YA
  - b. TIDAK
14. Jika hurufnya bukan L maka angkanya bukan 9.  
Hurufnya adalah L.  
Kesimpulan: Angkanya adalah 9.
  - a. YA
  - b. TIDAK
15. Jika hurufnya bukan J maka angkanya bukan 8.  
Angkanya bukan 8.  
Kesimpulan: Hurufnya bukan J.
  - a. YA
  - b. TIDAK
16. Jika hurufnya bukan V maka angkanya bukan 7.  
Angkanya 7.  
Kesimpulan: Hurufnya V.
  - a. YA
  - b. TIDAK

17. Jika hurufnya D maka angkanya 2.  
Hurufnya D.  
Kesimpulan: Angkanya 2.  
a. YA  
b. TIDAK
18. Jika hurufnya Q maka angkanya 1.  
Hurufnya K.  
Kesimpulan: Angkanya bukan 1.  
a. YA  
b. TIDAK
19. Jika hurufnya M maka angkanya 4.  
Angkanya 4.  
Kesimpulan: Hurufnya M.  
a. YA  
b. TIDAK
20. Jika hurufnya V maka angkanya 5.  
Angkanya 2.  
Kesimpulan: Hurufnya bukan V.  
a. YA  
b. TIDAK
21. Jika hurufnya S maka angkanya bukan 8.  
Hurufnya S.  
Kesimpulan: Angkanya bukan 8.  
a. YA  
b. TIDAK
22. Jika hurufnya B maka angkanya bukan 3.  
Hurufnya H.  
Kesimpulan: Angkanya 3.  
a. YA  
b. TIDAK
23. Jika hurufnya J maka angkanya bukan 2.  
Angkanya 7.  
Kesimpulan: Hurufnya J.  
a. YA  
b. TIDAK
24. Jika hurufnya U maka angkanya bukan 7.  
Angkanya 7.  
Kesimpulan: Hurufnya bukan U.  
a. YA  
b. TIDAK
25. Jika hurufnya bukan E maka angkanya 2.  
Hurufnya R.  
Kesimpulan: Angkanya 2.  
a. YA  
b. TIDAK

26. Jika hurufnya bukan A maka angkanya 6.  
Hurufnya A.  
Kesimpulan: Angkanya bukan 6.  
a. YA  
b. TIDAK
27. Jika hurufnya bukan C maka angkanya 9.  
Angkanya 9.  
Kesimpulan: Hurufnya bukan C.  
a. YA  
b. TIDAK
28. Jika hurufnya bukan N maka angkanya 3.  
Angkanya 5.  
Kesimpulan: Hurufnya N.  
a. YA  
b. TIDAK
29. Jika hurufnya bukan A maka angkanya bukan 1.  
Hurufnya N.  
Kesimpulan: Angkanya bukan 1.  
a. YA  
b. TIDAK
30. Jika hurufnya bukan C maka angkanya bukan 2.  
Hurufnya C.  
Kesimpulan: Angkanya 2.  
a. YA  
b. TIDAK
31. Jika hurufnya bukan W maka angkanya bukan 8.  
Angkanya 3.  
Kesimpulan: Hurufnya bukan W.  
a. YA  
b. TIDAK
32. Jika hurufnya bukan K maka angkanya bukan 1.  
Angkanya adalah 1.  
Kesimpulan: Hurufnya adalah K.  
a. YA  
b. TIDAK

### **3. LAMPIRAN C**

#### **3.1 Tugas Silogisme Bias Keyakinan**

Pada halaman berikut adalah tugas silogisme bias keyakinan dan instruksinya seperti yang disajikan kepada peserta. Tugas dibagi menjadi dua bagian seperti yang dijelaskan pada Bab 5. Bagian 1 adalah pertanyaan 1-12 dan bagian 2 adalah pertanyaan 13-24.

#### **3.2 Bernalar dengan informasi**

***Silakan baca petunjuk berikut dengan seksama.***

Dalam masalah berikut, Anda akan diberikan dua premis, yang harus Anda asumsikan benar. Sebuah kesimpulan dari premis kemudian berikut. Anda harus memutuskan apakah

kesimpulan mengikuti secara logis dari premis atau tidak. Anda harus menganggap bahwa semua premis itu benar dan membatasi diri Anda hanya pada informasi yang terkandung di dalam premis tersebut. Ini sangat penting. Putuskan apakah kesimpulannya mengikuti secara logis dari premis, dengan asumsi premis itu benar, dan centang jawaban Anda.

Jika menurut Anda kesimpulannya mengikuti secara logis, dengan asumsi bahwa premis-premisnya benar, silakan centang YA, jika tidak centang TIDAK. Jawab semua pertanyaan.

**Bagian 1:**

1. Premis: Semua hal yang dihisap baik untuk kesehatan.  
Rokok dihisap.  
Kesimpulan: Rokok baik untuk kesehatan.  
a. YA  
b. TIDAK
2. Premis: Segala sesuatu yang terbuat dari kayu dapat digunakan sebagai bahan bakar.  
Bensin tidak terbuat dari kayu.  
Kesimpulan: Bensin tidak dapat digunakan sebagai bahan bakar.  
a. YA  
b. TIDAK
3. Premis: Semua lapitar memakai pakaian.  
Podip memakai pakaian.  
Kesimpulan: Podip adalah lapitar.  
a. YA  
b. TIDAK
4. Premis: Semua kacang bisa dimakan.  
Batu tidak bisa dimakan.  
Kesimpulan: Batu bukanlah kacang.  
a. YA  
b. TIDAK
5. Premis: Semua orang miskin menganggur.  
Rockefeller tidak miskin.  
Kesimpulan: Rockefeller tidak menganggur.  
a. YA  
b. TIDAK
6. Premis: Semua senjata berbahaya.  
Ular derik berbahaya.  
Kesimpulan: Ular derik adalah senjata.  
a. YA  
b. TIDAK
7. Premis: Segala sesuatu dengan empat kaki berbahaya.  
Anjing tidak berbahaya.  
Kesimpulan: Anjing tidak memiliki empat kaki.  
a. YA  
b. TIDAK
8. Premis: Semua ramadions rasanya enak.  
Gumthorps adalah ramadion.  
Kesimpulan: Gumthorps rasanya enak.

- a. YA
  - b. TIDAK
9. Premis: Semua makhluk hidup membutuhkan air.  
Mawar membutuhkan air.  
Kesimpulan: Mawar adalah makhluk hidup.
- a. YA
  - b. TIDAK
10. Premis: Semua selacians memiliki gigi yang tajam.  
Snorlup tidak memiliki gigi yang tajam.  
Kesimpulan: Snorlup bukanlah selasian.
- a. YA
  - b. TIDAK
11. Premis: Semua ikan bisa berenang.  
Tuna adalah ikan.  
Kesimpulan: Tuna bisa berenang.
- a. YA
  - b. TIDAK
12. Premis: Semua hudon ganas.  
Wapet bukan hudon.  
Kesimpulan: Wampets tidak ganas.
- a. YA
  - b. TIDAK

**Bagian 2:**

13. Premis: Semua opprobines menggunakan listrik.  
Jamtop berjalan dengan listrik.  
Kesimpulan: Jamtop adalah opprobines.
- a. YA
  - b. TIDAK
14. Premis: Semua yang hidup minum air.  
Televisi tidak minum air.  
Kesimpulan: Televisi tidak hidup.
- a. YA
  - b. TIDAK
15. Premis: Semua kelelawar memiliki sayap.  
Elang bukan kelelawar.  
Kesimpulan: Elang tidak memiliki sayap.
- a. YA
  - b. TIDAK
16. Premis: Semua mamalia berjalan.  
Paus adalah mamalia.  
Kesimpulan: Paus berjalan.
- a. YA
  - b. TIDAK
17. Premis: Semua benda besar membutuhkan oksigen.  
Tikus membutuhkan oksigen.

- Kesimpulan: Tikus adalah makhluk besar.
- a. YA
  - b. TIDAK
18. Premis: Semua negara Afrika panas.  
Kanada bukan negara Afrika.  
Kesimpulan: Kanada tidak panas.
- a. YA
  - b. TIDAK
19. Premis: Semua hal yang bergerak menyukai air.  
Kucing tidak suka air.  
Kesimpulan: Kucing tidak bergerak.
- a. YA
  - b. TIDAK
20. Premis: Semua tupper bertelur.  
Sampet adalah tupper.  
Kesimpulan: Sampet bertelur.
- a. YA
  - b. TIDAK
21. Premis: Segala sesuatu yang memiliki motor membutuhkan oli.  
Mobil butuh oli.  
Kesimpulan: Mobil memiliki motor.
- a. YA
  - b. TIDAK
22. Premis: Semua snappers berjalan cepat.  
Alcoma tidak berjalan cepat.  
Kesimpulan: Alcoma bukanlah snappers.
- a. YA
  - b. TIDAK
23. Premis: Semua burung memiliki bulu.  
Robin adalah burung.  
Kesimpulan: Robin memiliki bulu.
- a. YA
  - b. TIDAK
24. Premis: Semua argomeles baik.  
Magsum bukan argomel.  
Kesimpulan: Magsum tidak baik.
- a. YA
  - b. TIDAK

#### 4. LAMPIRAN D

##### 4.1 Kebutuhan Skala Kognisi

###### *Gaya bernalar*

Berilah nilai pada pernyataan-pernyataan berikut menurut skala ini:

1 - 2 = Sangat tidak Setuju

3 - 4 = Kurang Setuju

*ketrampilan Penalaran Deduktif (Dr. Agus Wibowo)*

- 5 – 6 = Antara Setuju dan Tidak Setuju  
 7 – 8 = Cukup Setuju  
 9 = Sangat Setuju

Saya lebih suka masalah yang rumit daripada yang sederhana.	1	2	3	4	5	6	7	8	9
Saya suka memiliki tanggung jawab menangani situasi yang membutuhkan banyak penalaran.	1	2	3	4	5	6	7	8	9
Bernalar bukanlah ide saya untuk bersenang-senang.	1	2	3	4	5	6	7	8	9
Saya lebih suka melakukan sesuatu yang membutuhkan sedikit penalaran daripada sesuatu yang pasti akan menantang kemampuan bernalar saya.	1	2	3	4	5	6	7	8	9
Saya mencoba mengantisipasi dan menghindari situasi di mana ada kemungkinan saya harus memikirkan sesuatu secara mendalam.	1	2	3	4	5	6	7	8	9
Saya menemukan kepuasan dalam berunding dengan keras dan berjam-jam.	1	2	3	4	5	6	7	8	9
Saya hanya bernalar sekeras yang saya harus.	1	2	3	4	5	6	7	8	9
Saya lebih suka memikirkan proyek harian kecil daripada proyek jangka panjang.	1	2	3	4	5	6	7	8	9
Saya menyukai tugas yang membutuhkan sedikit penalaran setelah saya mempelajarinya.	1	2	3	4	5	6	7	8	9
Gagasan mengandalkan penalaran untuk mencapai puncak menarik bagi saya.	1	2	3	4	5	6	7	8	9
Saya sangat menikmati tugas yang melibatkan menemukan solusi baru untuk masalah.	1	2	3	4	5	6	7	8	9
Mempelajari cara-cara baru untuk bernalar tidak terlalu menggairahkan saya.	1	2	3	4	5	6	7	8	9
Saya lebih suka hidup saya dipenuhi dengan teka-teki yang harus saya pecahkan.	1	2	3	4	5	6	7	8	9
Gagasan bernalar abstrak menarik bagi saya.	1	2	3	4	5	6	7	8	9
Saya lebih suka tugas yang intelektual, sulit, dan penting daripada tugas yang agak penting tetapi tidak memerlukan banyak penalaran.	1	2	3	4	5	6	7	8	9
Saya merasa lega daripada puas setelah menyelesaikan tugas yang membutuhkan banyak usaha mental.	1	2	3	4	5	6	7	8	9
Cukup bagi saya bahwa sesuatu menyelesaikan pekerjaan; Saya tidak peduli bagaimana atau mengapa itu bekerja.	1	2	3	4	5	6	7	8	9
Saya biasanya akhirnya berunding tentang masalah bahkan ketika mereka tidak mempengaruhi saya secara pribadi.	1	2	3	4	5	6	7	8	9

## 5. LAMPIRAN E

### 5.1 Tes Matematika tingkat AS

Pemeriksaan manipulasi matematika yang diberikan kepada mahasiswa tingkat AS di Bab 5 disajikan di bawah ini bersama dengan instruksi. Soal 1-12 diambil dari subtes Perhitungan Woodcock-Johnson III. Sembilan dari mereka telah menunjukkan akurasi rata-rata kurang dari 55% dan berkorelasi dengan kinerja di seluruh tes di 0,86 dalam dataset sebelumnya dengan mahasiswa sarjana campuran (Inggris, Attridge et al., 2011) dan tiga sisanya diambil dari rentang yang lebih rendah untuk mencegah efek lantai pada mahasiswa bahasa. Pertanyaan 13-15 adalah pertanyaan yang paling sulit pada tes diagnostik Universitas Stekom untuk sarjana matematika baru berdasarkan kinerja pada tahun 2008 dan 2009. Tes diagnostik dirancang untuk menilai kemampuan mahasiswa dengan materi yang tercakup dalam matematika tingkat AS, dan tiga item ini disertakan untuk mencegah efek langit-langit pada mahasiswa matematika pada titik waktu kedua sambil memastikan bahwa konten tidak dikembangkan secara tidak tepat. Pertanyaan disajikan dalam urutan yang dimaksudkan untuk menjadi progresif.

### 5.2 Perhitungan

Beberapa pertanyaan di bagian ini sangat sulit, dan beberapa bahkan tidak akan pernah Anda temui sebelumnya. Kami tidak mengharapkan Anda untuk dapat mencoba semua pertanyaan, tolong jawab sebanyak yang Anda bisa. Jika Anda buntu, jangan khawatir, lanjutkan saja ke bagian berikutnya.

1.  $9 + 7 =$
2.  $8 \times 5 =$
3.  $48 - 19 =$
4.  $1.05 \times 0.2 =$
5.  $\left(\frac{4b}{3y}\right)\left(\frac{-4y}{12b^2}\right) =$
6.  $\log_b 81 = 4$   
b =
7.  $f(x) = 6x^3$   
 $f'(x) =$
8.  $\cos\theta = \frac{\sqrt{3}}{2}$
9.  $\begin{vmatrix} 8 & 2 \\ -4 & 1 \end{vmatrix}$
10.  $\int_1^3 3x^2 dx$
11.  $\tan\theta = 1$   
 $\sin\theta$
12.  $\int xe^x dx$  is
  - a.  $x^2e^x + c$
  - b.  $xe^x - e^x + c$
  - c.  $\frac{x^2}{2}e^x + c$
  - d.  $\frac{x^2}{2}e^{x+1} + c$
13. Jika 238
  - a.  $-\frac{1}{5(5x-7)} + c$



- b.  $\frac{1}{5x-7} + c$   
 c.  $-\frac{1}{5x-7} + c$   
 d.  $-\frac{5}{5x-7} + c$

## 6. LAMPIRAN F

### 6.1 Tugas Inferensi Kondisional Tematik

Di sini, tugas inferensi kondisional tematik yang diberikan kepada mahasiswa sarjana di Bab 5 disajikan dengan instruksinya. Tidak ada batasan waktu.

### 6.2 Logika dunia nyata

***Silakan baca petunjuk berikut dengan seksama:***

Studi ini berkaitan dengan kemampuan orang untuk bernalar secara logis dengan kalimat sehari-hari. Anda akan disajikan dengan total 16 masalah pada halaman berikut. Dalam setiap kasus Anda diberikan dua pernyataan bersama dengan kesimpulan yang mungkin atau mungkin tidak mengikuti dari pernyataan-pernyataan ini.

Tugas Anda dalam setiap kasus adalah mengevaluasi logika kalimat, terlepas dari isinya, dan memutuskan apakah kesimpulan harus mengikuti atau tidak dari pernyataan. Sebuah kesimpulan diperlukan jika itu harus benar, mengingat bahwa pernyataan itu benar.

Masalahnya terlihat seperti ini:

*Anggapan berikut ini benar:*

Jika bahannya kayu, maka itu keras.

*Mengingat bahwa premis berikut ini juga benar:*

Bahannya adalah kayu.

*apakah perlu bahwa:*

Bahannya keras.

- a. YA  
 b. TIDAK

Dalam setiap kasus, Anda harus mengasumsikan bahwa aturan tersebut berlaku dan menggabungkannya dengan informasi dalam pernyataan kedua untuk memutuskan apakah kesimpulan harus mengikuti.

Jika menurut Anda kesimpulannya harus mengikuti, silakan centang kotak YA, jika tidak centang kotak TIDAK. Silakan selesaikan masalah secara berurutan dan pastikan Anda tidak melewatkannya. Jangan kembali ke masalah setelah Anda selesai dan pindah ke yang lain.

Jika menurut Anda kesimpulannya harus mengikuti, silakan centang YA, jika tidak centang TIDAK. Jangan kembali ke masalah setelah Anda selesai dan pindah ke yang lain. Jawab semua pertanyaan.

1. *Asumsikan hal berikut ini benar:*

Jika harga minyak terus naik maka harga bensin Indonesia akan naik.

*Mengingat bahwa premis berikut ini juga benar:*

Harga minyak terus naik.

*apakah perlu bahwa:*

harga bensin Indonesia naik.

- a. YA  
 b. TIDAK

2. *Anggapan berikut ini benar:*

- Jika kepemilikan mobil meningkat maka kemacetan lalu lintas akan semakin parah.  
*Mengingat bahwa premis berikut ini juga benar:*  
 Kemacetan lalu lintas tidak bertambah parah.  
*apakah perlu:*  
 Kepemilikan mobil tidak bertambah.
- a. YA
  - b. TIDAK
3. *Anggapan berikut ini benar:*  
 Jika lebih banyak orang menggunakan krim pelindung matahari maka kasus kanker kulit akan berkurang.  
*Mengingat premis berikut ini juga benar:*  
 Kasus kanker kulit tidak berkurang.  
*apakah perlu bahwa:*  
 Lebih banyak orang tidak menggunakan krim pelindung matahari.
- a. YA
  - b. TIDAK
4. *Asumsikan hal berikut ini benar:*  
 Jika Sony merilis PlayStation 4 maka keuntungan perusahaan mereka akan meningkat.  
*Mengingat bahwa premis berikut ini juga benar:*  
 Sony merilis PlayStation 4.  
*apakah perlu bahwa:*  
 Laba perusahaan meningkat.
- a. YA
  - b. TIDAK
5. *Asumsikan hal berikut ini benar:*  
 Jika harga minyak terus naik maka harga bensin Indonesia akan naik.  
*Mengingat bahwa premis berikut ini juga benar:*  
 Harga minyak tidak terus naik.  
*apakah perlu bahwa:*  
 harga bensin Indonesia tidak naik.
- a. YA
  - b. TIDAK
6. *Anggapan berikut ini benar:*  
 Jika kepemilikan mobil meningkat maka kemacetan lalu lintas akan semakin parah.  
*Mengingat bahwa premis berikut ini juga benar:*  
 Kemacetan lalu lintas semakin parah.  
*apakah perlu bahwa:*  
 Kepemilikan mobil meningkat.
- a. YA
  - b. TIDAK
7. *Anggapan berikut ini benar:*  
 Jika lebih banyak orang menggunakan krim pelindung matahari maka kasus kanker kulit akan berkurang.  
*Mengingat bahwa premis berikut ini juga benar:*  
 Kasus kanker kulit berkurang.

- apakah perlu bahwa:*  
 Lebih banyak orang menggunakan krim pelindung matahari.
- a. YA
  - b. TIDAK
8. *Asumsikan hal berikut ini benar:*  
 Jika Sony merilis PlayStation 4 maka keuntungan perusahaan mereka akan meningkat.  
*Mengingat premis berikut ini juga benar:*  
 Sony tidak merilis PlayStation 4.  
*apakah perlu bahwa:*  
 Laba perusahaan tidak naik.
- a. YA
  - b. TIDAK
9. *Asumsikan hal berikut ini benar:*  
 Jika lebih banyak rumah baru dibangun maka jumlah tunawisma akan meningkat.  
*Mengingat bahwa premis berikut ini juga benar:*  
 Lebih banyak rumah baru dibangun.  
*apakah perlu bahwa:*  
 Jumlah tunawisma meningkat.
- a. YA
  - b. TIDAK
10. *Anggapan berikut ini benar:*  
 Jika utang dunia ketiga dibatalkan maka kemiskinan dunia akan semakin parah.  
*Mengingat bahwa premis berikut ini juga benar:*  
 Kemiskinan dunia tidak memburuk.  
*apakah perlu bahwa:*  
 Hutang dunia ketiga tidak dibatalkan.
- a. YA
  - b. TIDAK
11. *Anggapan berikut ini benar:*  
 Jika makanan cepat saji dikenai pajak maka obesitas pada anak akan meningkat.  
*Mengingat bahwa premis berikut ini juga benar:*  
 Obesitas pada masa kanak-kanak tidak meningkat.  
*apakah perlu bahwa:*  
 Makanan cepat saji tidak dikenakan pajak.
- a. YA
  - b. TIDAK
12. *Asumsikan hal berikut ini benar:*  
 Jika undang-undang karantina Indonesia diperkuat maka virus covid19 akan menyebar ke Malaysia.  
*Mengingat bahwa premis berikut ini juga benar:*  
 undang-undang karantina Indonesia diperkuat.  
*apakah perlu bahwa:*  
 Rabies menyebar ke Indonesia.
- a. YA
  - b. TIDAK

## 7. LAMPIRAN G

### 7.1 Tes Matematika Sarjana

Tes matematika yang diberikan kepada mahasiswa S1 pada Bab 5 terdiri dari 11 soal. Tujuh di antaranya diambil dari subtes Perhitungan Woodcock-Johnson III, dua adalah pertanyaan paling sulit pada tes diagnostik Universitas Stekom untuk sarjana matematika baru berdasarkan kinerja pada tahun 2008 dan 2009, dan dua yang terakhir didasarkan pada gelar matematika tahun pertama. Silabus.

### 7.2 Perhitungan

Beberapa pertanyaan di bagian ini sangat sulit, dan beberapa bahkan tidak akan pernah Anda temui sebelumnya. Kami tidak mengharapkan Anda untuk dapat mencoba semua pertanyaan, tolong jawab sebanyak yang Anda bisa. Jika Anda buntu, jangan khawatir, lanjutkan saja ke bagian berikutnya. Anda boleh membuat catatan tapi tolong jangan gunakan kalkulator.

1.  $8 \times 5 =$
2.  $48 - 19 =$
3.  $1.05 \times 0.2 =$
4.  $\left(\frac{4b}{3y}\right)\left(\frac{-4y}{12b^2}\right) =$
5.  $f(x) = 6x^3$   
 $f'(x) =$
6.  $2y = 6x + 8$   
Slope =  
y-intersept =
7.  $\begin{vmatrix} 8 & 2 \\ -4 & 1 \end{vmatrix}$
8.  $\int_1^3 3x^2 dx$
9. Jika 238
  - a.  $-\frac{1}{5(5x-7)} + c$
  - b.  $\frac{1}{5x-7} + c$
  - c.  $-\frac{1}{5x-7} + c$
  - d.  $-\frac{5}{5x-7} + c$
10. Saat mengekspresikan  $\frac{x}{(x+1)^2(x^2+2)}$

## DAFTAR PUSTAKA

- Alter, A. L., Oppenheimer, D. M., Epley, N. & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, 136, 569-576.
- American Council on Education. (1953). *Test of critical thinking: Instructor's manual*. Washington, D.C: American Council on Education.
- Anderson, A. R. & Belnap, N. D. (1975). *Entailment: The logic of relevance and necessity*. Princeton, NJ: Princeton University Press.
- Ardila, A., Pineda, D. & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, 15, 31-36.
- Arffa, S. (2007). The relationship of intelligence to executive function and nonexecutive function measures in a sample of average, above average, and gifted youth. *Archives of Clinical Neuropsychology*, 22, 969-978.
- Aronson, J., Lustina, M. J., Good, C. & Keough, K. (1999). When white men can't do math: necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, 35, 29-46.
- Ashenfelter, O. & Rouse, C. (1999, January). *Schooling, intelligence, and income in America: Cracks in the Bell curve* (Working Paper No. 6902). Princeton: National Bureau of Economic Research. Available from <http://www.nber.org/papers/w6902>
- Baddeley, A. & Hitch, G. (1974). The psychology of learning and motivation: advances in research and theory. In G. H. Bower (Ed.), (Vol. 8, p. 47-89). New York: Academic Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon.
- anich, M. T. (2009). Executive function: the search for an integrated account. *Current Directions in Psychological Science*, 18, 89-94.
- Baron, J. & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, 54, 569-579.
- Beth, E. W. & Piaget, J. (1966). *Mathematical epistemology and psychology*. Dordrecht: D. Reidel.
- Bors, D. A. & Vigneau, F. (2003). The effect of practice on Raven's Advanced Progressive Matrices. *Learning and Individual Differences*, 13, 291-312.
- Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and the standard logic. *Psychological Review*, 85, 1-21.
- Brainerd, C. J. & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, 11, 164-169.
- Bramall, S. (2000). *Rethinking the place of mathematical knowledge in the curriculum*.
- Bramall, S. & J. White (Eds.), *Why learn maths?* London: Institute of Education University of London.

- Bramall, S. & White, J. (Eds.). (2000). *Why learn maths*. London: Institute of Education University of London.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A. & Jarvis, W. B. G. (1996). Individual difference in cognitive motivation: the life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197-253.
- Cacioppo, J. T., Petty, R. E. & Kao, C. F. (1984). The efficient assessment of Need for Cognition. *Journal of Personality Assessment*, 48, 306-307.
- Chen, S. & Chaiken, S. (1999). Dual process theories in social psychology. In S. Chaiken & Y. Trope (Eds.), (p. 73-96). New York: Guilford.
- Cheng, P. W. & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391-416.
- Cheng, P. W., Holyoak, K. J., Nisbett, R. E. & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, 18, 293-328.
- Christensen, L. B. (2000). *Experimental methodology* (Seventh ed.). Allyn and Bacon.
- Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? studies with the Wason selection task. *Cognition*, 31, 187-316.
- Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange  
J. H. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind*. Oxford: Oxford University Press.
- Cummins, D. D. (1995). Naive theories and causal deduction. *Memory & Cognition*, 23, 646-658.
- De Neys, W. (2006). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *The Quarterly Journal of Experimental Psychology*, 59, 1070-1100.
- De Neys, W., Schaeken, W. & d'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: retrieval and inhibition of stored counterexamples. *Thinking & Reasoning*, 11, 349-381.
- Deary, I. J. (2001). *Intelligence: a very short introduction*. Oxford: Oxford University Press.
- Deary, I. J. (2008). Why do intelligent people live longer? *Nature*, 456, 175-176.
- Diener, E. & Crandall, R. (1978). *Ethics in social and behavioural research*. Chicago: University of Chocago Press.
- Dolton, P. J. & Vignoles, A. (2002). The return on post-compulsory school mathematics study. *Economica*, 69, 113-141.
- Elias, S. M. & Loomis, R. J. (2002). Utilizing Need for Cognition and percieved self-efficacy to predict academic performance. *Journal of Applied Social Psychology*, 32, 1687-1702.
- Erwin, T. D. (1981). *Manual for the scale of intellectual development*. Harrisonburg, VA: Developmental Analytics. ESRC. (2009). *ESRC strategic plan 2009-2014*. Swindon: ESRC.
- Ethics Committe of the British Psychological Society. (2009). *Code of Ethics and Conduct*. Leicester: The British Psychological Society.

- Evans, J. St. B. T. (1977). Linguistic factors in reasoning. *Quarterly Journal of Experimental Psychology*, 29, 297-306.
- Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology*, 75, 451-468.
- Evans, J. St. B. T. (1993). The mental model theory of conditional reasoning: critical appraisal and revision. *Cognition*, 48, 1-20.
- Evans, J. St. B. T. (1998). Matching bias in conditional reasoning: do we understand it after 25 years? *Thinking and Reasoning*, 4, 45-82.
- Evans, J. St. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7, 454-459.
- Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, 13, 378-395.
- Evans, J. St. B. T. (2007). *Hypothetical thinking: dual processes in reasoning and judgement*. Hove: Psychology Press.
- Evans, J. St. B. T. (2011). Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review*, 31, 86102.
- Evans, J. St. B. T., Barston, J. L. & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, 11, 295306.
- Evans, J. St. B. T., Clibbens, J. & Rood, B. (1995). Bias in conditional inference: implications for mental models and mental logic. *The Quarterly Journal Of Experimental Psychology*, 48A, 644-670.
- Evans, J. St. B. T. & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking and Reasoning*, 11, 382-389.
- Evans, J. St. B. T. & Handley, S. J. (1999). The role of negation on conditional inference. *The Quarterly Journal of Experimental Psychology*, 52A, 739769.
- Evans, J. St. B. T., Handley, S. J. & Bacon, A. M. (2009). Reasoning under time pressure: A study of causal conditional inference. *Experimental Psychology*, 56, 77-83.
- Evans, J. St. B. T., Handley, S. J., Neilens, H. & Over, D. (2010). The influence of cognitive ability and instructional set of causal conditional inference. *The Quarterly Journal of Experimental Psychology*, 63, 892-909.
- Evans, J. St. B. T., Handley, S. J., Neilens, H. & Over, D. E. (2007). Thinking about conditionals: a study of individual differences. *Memory & Cognition*, 35, 1772-1784.
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Executive Summary "The Delphi Report". Millbrae, CA: California Academic Press.
- Farsides, T. & Woodfield, R. (2003). Individual differences and undergraduate academic success: the roles of personality, intelligence, and application. *Personality and Individual Differences*, 34, 1225-1243.

- Fong, G. T., Krantz, D. H. & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, 18, 253-292.
- Fong, G. T. & Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology: General*, 120, 34-45.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19, 25-42.
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C. & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, 17, 172-179.
- Frisch, D. (1993). Reasons for framing effects. *Organizational Behaviour and Human Decision Processes*, 54, 399-429.
- Garcia-Madruga, J. A., Gutierrez, F., Carriedo, N., Luzon, J. M. & Vila, J. O. (2007). Mental models in propositional reasoning and working memory's central executive. *Thinking & Reasoning*, 13, 370-393.
- George, D. & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.)*. Boston: Allyn & Bacon.
- Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451-482.
- Gilhooly, K. J. & Fioratou, E. (2009). Executive functions in insight versus noninsight problem solving: An individual differences approach. *Thinking & Reasoning*, 15, 355-376.
- Gillard, E. (2009). *Dual processes in the psychology of mathematics education and beyond*. Unpublished doctoral dissertation, Katholieke Universiteit Leuven.
- Gillard, E., Van Dooren, W., Schaeken, W. & Verschaffel, L. (2009a). Processing time evidence for a default-interventionist model of probability judgements. In *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (p. 1792-1797). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gillard, E., Van Dooren, W., Schaeken, W. & Verschaffel, L. (2009b). Proportional reasoning as a heuristic-based process. *Experimental Psychology*, 56, 92-99.
- Giroto, V., Kimmelmeir, M., Sperber, D. & van der Henst, J. B. (2001). Inept reasoners or pragmatic virtuosos? relevance and the deontic selection task. *Cognition*, 81, 69-76.
- Giroto, V. & Legrenzi, P. (1989). Mental representation and hypothetico-deductive reasoning: The case of the THOG problem. *Psychological Research*, 51, 129-135.
- Giroto, V. & Legrenzi, P. (1993). Naming the parents of THOG: mental representation and reasoning. *Quarterly Journal of Experimental Psychology*, 46A, 701-713.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M. & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgement. *Science*, 293, 2105-2108.
- Greenwald, A. G. & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem and stereotypes. *Psychological Review*, 102, 4-27.



- Handley, S. J., Capon, A., Beveridge, M., Dennis, I. & Evans, J. St. B. T. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking & Reasoning*, 10, 175-195.
- Handley, S. J., Capon, A., Copp, C. & Harper, C. (2002). Conditional reasoning and the tower of hanoi: the role of spatial and verbal working memory. *British Journal of Psychology*, 93, 501-518.
- Harman, G. (1995). Thinking: invitation to cognitive science. In E. E. Smith & D. N. Osherson (Eds.), (Vol. 3, p. 175-211). Cambridge: MIT Press. Heim, A. (1969). AH5 group test of intelligence. London: National Foundation of Educational Research.
- Heiman, G. W. (2002). *Research methods in psychology*. Boston, MA: Houghton Mifflin.
- Heit, E. & Rotello, C. M. (2010). Relations between inductive reasoning and deductive reasoning. *Journal of Experimental Psychology: Learning, memory and Cognition*, 36, 805-812.
- Hergenhahn, B. R. & Olson, M. H. (2004). *An introduction to theories of learning (7th ed.)*. New Jersey: Prentice Hall.
- Houston, K. (2009). *How to think like a mathematician: a companion to undergraduate mathematics*. Cambridge: Cambridge University Press.
- Hoyles, C. & Kuchemann, D. (2002). Students' understanding of logical implication. *Educational Studies in Mathematics*, 51, 193-223.
- Huckstep, P. (2000). Mathematics as a vehicle for 'mental training'. In S. Bramall & J. White (Eds.), *Why learn maths?* London: Institute of Education University of London.
- Inglis, M. (2012). Views on the theory of formal discipline. Unpublished manuscript.
- Inglis, M., Attridge, N., Batchelor, S. & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin & Review*, 18, 1222-1229.
- Inglis, M., Palipana, A., Trenholm, S. & Ward, J. (2011). Individual differences in students' use of optimal learning resources. *Journal of Computer Assisted Learning*, 27, 490-502.
- Inglis, M. & Simpson, A. (2004). Mathematicians and the selection task. In M. Johnsen & A. B. Fuglestad (Eds.), *Proceedings of the 28th conference of the international group for the psychology of mathematics education (Vol. 3, p. 89-96)*.
- Inglis, M. & Simpson, A. (2008). Conditional inference and advanced mathematical study. *Educational Studies in Mathematics*, 67, 187-204.
- Inglis, M. & Simpson, A. (2009a). Conditional inference and advanced mathematical study: Further evidence. *Educational Studies in Mathematics*, 72, 185-198.
- Inglis, M. & Simpson, A. (2009b). The defective and material conditionals in mathematics: does it matter? In *33rd Conference of the International Group for the Psychology of Mathematics Education, PME 33 (p. 225232)*.
- Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

- Jensen, A. R. (1998). *The g factor: the science of mental ability*. Westport, CT: Greenwood.
- Johnson-Laird, P. N. (2008). *How we reason*. Oxford: Oxford University Press. Johnson-Laird, P. N. & Byrne, R. M
- Johnson-Laird, P. N. & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109, 646-678.
- Johnson-Laird, P. N., Byrne, R. M. J. & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, 99, 418-439.
- Johnson-Laird, P. N., Legrenzi, P. & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, 63, 395-400.
- Johnson-Laird, P. N. & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, 1, 134-148.
- Jones, W., Russell, D. & Nickel, T. (1977). Belief in the paranormal scale: an objective instrument to measure belief in magical phenomena and causes. *JSAS Catalog of Selected Documents in Psychology*, 7, 1-32.
- Judge, T. A., Higgins, C. A., Thoresen, C. J. & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52, 621-652.
- Kahneman, D. (1991). Judgment and decision making: A personal view. *Psychological Science*, 2, 142-145.
- Kahneman, D. & Tversky, A. (1972). On prediction and judgement. *Oregon Research Institute Bulletin*, 12, 4.
- Kilpatrick, J. (1983). Research problems in mathematics education. *For the Learning of Mathematics*, 4(1), 45-46.
- Kosonen, P. & Winne, P. H. (1995). Effects of teaching statistical laws on reasoning about everyday problems. *Journal of Educational Psychology*, 87, 33-46.
- LaBerge, D. & Samuels, S. K. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.
- Landsberger, H. A. (1958). *Hawthorne revisited*. Ithaca, NY: Cornell University.
- Larrick, R. P., Nisbett, R. E. & Morgan, J. N. (1993). Who uses the cost-benefit rules of choice? implications for the normative status of microeconomic theory. *Organizational Behaviour and Human Decision Processes*, 56, 331-347.
- Larsen, L., Hartmann, P. & Nyborg, H. (2008). The stability of general intelligence from adulthood to middle-age. *Intelligence*, 36, 29-34.
- Lawson, D. (1997). What can we expect from a level mathematics students? *Teaching Mathematics and its Applications*, 16, 151-156.
- Lawson, D. (2003). Changes in student entry competences 1991-2001. *Teaching Mathematics and its Applications*, 22, 171-175.

- Lehman, D. R., Lempert, R. O. & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist*, 43, 431-442.
- Lehman, D. R. & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Developmental Psychology*, 26, 952-960.
- Lehmann, I. J. (1963). Changes in critical thinking, attitudes, and values from freshman to senior years. *Journal of Educational Psychology*, 54(6), 305-315.
- Lewin, I. (1982). Driver training: a perceptual-motor skill approach. *Ergonomics*, 25, 917-924.
- Locke, J. (1971/1706). *Conduct of the Understanding*. New York: Burt Franklin.
- Locurto, C. (1990). The malleability of IQ as judged from adoption studies. *Intelligence*, 14, 275-292.
- Manktelow, K. (1999). *Reasoning and thinking*. Hove, East Sussex: Psychology Press.
- Markovits, H. (1985). Incorrect conditional reasoning among adults: competence or performance? *British Journal of Psychology*, 76, 241-247.
- Markovits, H. & Doyon, C. (2004). Information processing and reasoning with premises that are empirically false: interference, working memory, and processing speed. *Memory & Cognition*, 32, 592-601.
- Markovits, H. & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory and Cognition*, 17, 11-17.
- Miller, P. H. (2011). *Theories of developmental psychology (Vol. 4th)*. New York, NY: Worth.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H. & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology*, 41, 49-100.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38, 379-387.
- Moutier, S., Angeard, N. & Houdé, O. (2002). Deductive reasoning and matching-bias inhibition training: evidence from a debiasing paradigm. *Thinking & Reasoning*, 8, 205-224.
- Nair, K. U. & Ramnarayan, S. (2000). Individual differences in Need for Cognition and complex problem solving. *Journal of Research in Personality*, 34, 305-328.
- National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington, D.C: National Academy Press.
- Neilens, H. L., Handley, S. J. & Newstead, S. E. (2009). Effects of training and instruction on analytic and belief-based reasoning processes. *Thinking and Reasoning*, 15, 37-68.
- Neisser, U. (1967). *Cognitive psychology*. New York, NY: Meredith.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J. et al. (1996). Intelligence: knowns and unknowns. *American Psychologist*, 51, 77-101.

- Newstead, S. E. (2000). Are there two different types of thinking? (Peer commentary on "Individual differences in reasoning: implications for the rationality debate?" by K. E. Stanovich and R. F. West). *Behavioural and Brain Sciences*, 23, 645-726.
- Newstead, S. E., Girotto, V. & Legrenzi, P. (1995). The THOG problem and its implications for human reasoning.
- In S. E. Newstead & J. St. B. T. Evans (Eds.), *Perspectives on thinking and reasoning. Essays in honour of Peter Wason*. Hove, UK: Lawrence Erlbaum Associates.
- Nisbett, R. E. (2009). *Can reasoning be taught?* Cambridge, MA: American Academy of Arts and Sciences.
- Nisbett, R. E., Fong, G. T., Lehman, D. R. & Cheng, P. W. (1987). Teaching reasoning. *Science*, 238, 625-631.
- Nisbett, R. E., Krantz, D. H., Jepson, C. & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, 90, 339-363.
- Novick, M. R. & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, 32, 1-13.
- Oakley, C. O. (1949). Mathematics. *The American Mathematical Monthly*, 56, 19.
- Oaksford, M., Chater, N. & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, memory and Cognition*, 26, 883-899.
- O'Brien, D. (2009). Human reasoning includes a mental logic. *Behavioural and Brain Sciences*, 32, 96-97.
- O'Brien, D. & Manfrinati, A. (2010). The mental logic theory of conditional propositions.
- In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: probability and logic in human thinking*. Oxford: Oxford University Press.
- Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S. et al. (2010). Putting brain training to the test. *Nature*, 465, 775-779.
- Pacini, R. & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio bias phenomenon. *Journal of Personality and Social Psychology*, 76, 972-987.
- Petty, R. E. & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Advances in experimental social psychology* (Vol. 19, p. 123-205). New York: Academic Press.
- Piaget, J. (1928). *Judgement and reasoning in the child*. London: Routledge.
- Piaget, J. (1960). *Logic and psychology*. New York: Basic Books. Piaget, J. (1970). *Piaget's theory*. In P. Mussen (Ed.), *Carmichael's manual of child psychology*. New York: Wiley. Plato. (2003/375B.C). *The Republic* (D. Lee, Ed.). London: Penguin.
- Pollard, P. & Evans, J. St. B. T. (1980). The influence of logic on conditional reasoning performance. *Quarterly Journal of Experimental Psychology*, 32, 605-624.

- Polya, G. (1954). *Induction and analogy in mathematics*. New Jersey: Princeton University Press.
- Pyszczynski, T., Greenberg, J. & Solomon, S. (1999). A dual-process model of defense against conscious and unconscious death-related thoughts: an extension of terror management theory. *Psychological Review*, 106, 835-845.
- Raven, J., Raven, J. C. & Court, J. H. (1998). *Manual for raven's advanced progressive matrices and vocabulary scales*. San Antonio: Pearson.
- Reeve, C. L. & Lam, H. (2005). The psychometric paradox of practice effects due to retesting: measurement invariance and stable ability estimates in the face of observed score changes. *Intelligence*, 33, 535-549.
- Reyna, V. F. & Brainerd, C. J. (1995). Fuzzy-trace theory: an interim synthesis. *Learning and Individual Differences*, 7, 1-75.
- Rips, L. (1989). The psychology of knights and knaves. *Cognition*, 31(2), 85-116. Rokeach, M. (1960). *The open and closed mind*. New York: Basic Books.
- Rönnlund, M. & Nilsson, L.-G. (2006). Adult life-span patterns in WAIS-R block design performance: cross-sectional versus longitudinal age gradients and relations to demographic factors. *Intelligence*, 34, 63-78.
- Ruston, J. P. & Jensen, A. R. (2010). The rise and fall of the Flynn Effect as a reason to expect a narrowing of the Black-White iq gap. *Intelligence*, 38, 213-219.
- Sá, W. C., West, R. F. & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalisable critical thinking skill. *Journal of Educational Psychology*, 91, 497-510.
- Sanz de Acedo Lizarraga, M. L., Sanz de Acedo Baquedano, M. T. & Soria Oliver, M. (2010). Psychological intervention in thinking skills with primary education students. *School Psychology International*, 31, 131-145.
- Schneider, W., Eschman, A. & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools, Inc.
- Schroyens, W., Schaeken, W. & d'Ydewalle, G. (2001). The processing of negations in conditional reasoning: A meta-analytical case study in mental model and/or mental logic theory. *Thinking and Reasoning*, 7, 121-172.
- Sfard, A. (1998). A mathematician's view of research in mathematics education: An interview with Shimson A. Amitsur. In A. Sierpiska & J. Kilpatrick (Eds.), *Mathematics education as a research domain: A search for identity* (Vol. 2, p. 445-458).
- Dordrecht: Kluwer. Shafir, E. (1994). Uncertainty and the difficulty of thinking through disjunctions. *Cognition*, 50, 403-430.
- Shih, M., Pittinsky, T. L. & Ambady, N. (1999). Stereotype susceptibility: identity salience and shifts in quantitative performance. *Psychological Science*, 10, 80-83.
- Skinner, B. F. (1938). *The behaviour of organisms*. New York: Appleton-Century-Crofts.

- Slovan, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22.
- Smith, A. (2004). *Making mathematics count: The report of Professor Adrian Smith's inquiry into post-14 mathematics education*. London: The Stationery Office.
- Spearman, C. (1927). *The abilities of of man*. New York: Macmillan.
- Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Sperber, D., Cara, F. & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, 57, 31-96.
- Sperber, D. & Girotto, V. (2002). Use of misuse of the selection task? Rejoinder to Fiddick, Cosmides, and Tooby. *Cognition*, 85, 277-290.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual difference in reasoning*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago: Univeristy of Chicago Press.
- Stanovich, K. E. (2009a). In two minds. In J. St. B. T. Evans & K. Frankish (Eds.), (p. 55-88). Oxford: Oxford University Press.
- Stanovich, K. E. (2009b). *What intelligence tests miss: The psychology of rational thought*. Yale: Yale University Press.
- Stanovich, K. E. & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: the cognitive correlates of print exposure. *Memory & Cognition*, 20, 51-68.
- Stanovich, K. E. & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, 89, 342-357.
- Stanovich, K. E. & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, 127, 161-188.
- Stanovich, K. E. & West, R. F. (2000). Individual differences in reasoning: implication for the rationality debate. *Behavioural and Brain Sciences*, 23, 645-726.
- Stanovich, K. E. & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Personality Processes and Individual Differences*, 94, 672-695.
- Stich, S. (1990). *The fragmentation of reason*. Cambridge, MA: MIT Press.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
- The British Psychological Society. (2010). *Code of human research ethics*. Leicester: The British Psychological Society.

- Thompson, V. A. (2009). Dual process theories: a metacognitive perspective. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: dual processes and beyond*. Oxford: Oxford University Press.
- Thompson, V. A. (2010). Towards a metacognitive dual process theory of conditional reasoning. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: probability and logic in human thinking*. Oxford: Oxford University Press.
- Thompson, V. A., Prowse Turner, J. A. & Pennycook, G. (2011). Intuition, reason and metacognition. *Cognitive Psychology*, 63, 107-140.
- Thorndike, E. L. (1924). Mental discipline in high school studies. *The Journal of Educational Psychology*, 15, 1-22.
- Thorndike, E. L. & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, 8, 247-261.
- Tobacyk, J. & Milford, G. (1983). Belief in paranormal phenomena. *Journal of Personality and Social Psychology*, 44, 1029-1037.
- Toms, M., Morris, N. & Ward, D. (1993). Working memory and conditional reasoning. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 46, 679-699.
- Toplak, M. E. & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, 94, 197-209.
- Toplak, M. E., West, R. F. & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39, 1275-1289.
- Troldhal, V. & Powell, F. (1965). A short-form dogmatism scale for use in field studies. *Social Forces*, 44, 211-215.
- Tversky, A. & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76, 105-110.
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453-458.
- Tversky, A. & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review*, 90, 293-315.
- Vamvakoussi, X., Van Dooren, W. & Verschaffel, L. (2012a). Educated adults are still affected by intuitions about the effect of arithmetical operation: evidence from a reaction-time

- study. *Educational Studies in Mathematics*, Advance online publication. DOI: 10.1007/s10649-012-9432-8.
- Vamvakoussi, X., Van Dooren, W. & Verschaffel, L. (2012b). Naturally biased? in search of reaction time evidence for a natural number bias in adults. *The Journal of Mathematical Behavior*, 31, 344-355.
- Van Breukelen, G. J. P. (2006). Ancova versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, 59, 920-925.
- Venet, M. & Markovits, H. (2001). Understanding uncertainty with abstract conditionals. *Merrill-Palmer Quarterly*, 47, 74-99.
- Verschueren, N., Schaeken, W. & d'Ydewalle, G. (2005). Everyday conditional reasoning: a working memory-dependent tradeoff between counterexample and likelihood use. *Memory & Cognition*, 33, 107-119.
- Walport, M. (2010). *Science and mathematics secondary education for the 21st century: Report of the Science and Learning Expert Group*. London: Crown.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology*. Harmondsworth: Penguin.
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20(3), 273-281.
- Wason, P. C. & Brooks, P. J. (1979). THOG: the anatomy of a problem. *Psychological Research*, 41, 79-90.
- Wason, P. C. & Johnson-Laird, P. N. (1972). *Psychology of reasoning: structure and content*. London: Batsford.
- Wason, P. C. & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, 63-71.
- West, R. F., Toplak, M. E. & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, 100, 930-941.
- White, E. E. (1936). A study of the possibility of improving habits of thought in school children by a training in logic. *British Journal of Educational Psychology*, 6, 267-273.
- Zepp, R. A. (1987). Logic in everyday language and in mathematics. *Chinese University Education Journal*, 15, 45-49.