

BAB III

METODOLOGI PENELITIAN

3.1 Jenis dan Sifat Penelitian

Penelitian ini menggunakan jenis kuantitatif sebagai upaya untuk mengevaluasi persepsi audiens terhadap dua versi *anchor AI tvOne*. Punch (2005) menyatakan data kuantitatif berbentuk angka-angka dan pengukuran merupakan proses untuk mengubah data menjadi angka-angka. Sugiyono (2013) juga menyatakan kuantitatif berlandaskan positivisme dengan tujuan untuk meneliti populasi menggunakan sampel yang bersifat statistik dengan tujuan untuk menguji hipotesis yang telah ditentukan.

Selebihnya, kuantitatif dalam penelitian ini bersifat komparatif kausal yang berdasarkan Werang (2020) perlu dibedakan dengan jenis korelasional. Menurut Werang (2020) penelitian komparatif kausal merupakan langkah lanjutan dari bentuk korelasional untuk menemukan hubungan sebab-akibat antara variabel. Selebihnya, Martella dalam Werang (2020) menyatakan bahwa penelitian komparatif kausal juga menggunakan lebih dari satu kelompok yang sesuai dengan penelitian ini yang membandingkan persepsi audiens terhadap *anchor AI tvOne* yang memiliki suara dan gerakan muka robotik, serta tidak terlihat untuk melakukan gestur tangan (Versi Lama), dibanding dengan *anchor AI tvOne* yang memiliki suara dan gerakan muka organik, serta terlihat untuk melakukan gestur tangan. itu, hasil dari tingkat antropomorfisme itu dikorelasikan dengan variabel dependen *intention to watch* dan *affective engagement*. Sebelumnya, riset-riset terdahulu terhadap *anchor AI* juga menggunakan semacam sifat komparatif seperti Kim et al. (2022) yang membandingkan kredibilitas *newscaster* dengan manusia. Adapun, Zhou et al. (2024) dan Xue et al. (2022) yang melakukan riset komparatif wujud dan sikap *news anchor AI* dan hubungannya dalam meningkatkan keinginan untuk berinteraksi dengan paparan informasi seperti mengomentari atau menyukai dan mengonsumsi berita oleh *news anchor AI*.

3.2 Metode Penelitian

Suwanda (2011) setuju dengan definisi eksperimen oleh Webster yaitu merupakan kata yang sama banding dengan percobaan atau pengamatan tertentu yang dirancang untuk menekankan atau membuktikan keadaan sebaliknya dari sesuatu yang belum pasti. Ia pun juga menambahkan definisi yakni eksperimen dirancang dengan kondisi-kondisi tertentu oleh peneliti. Menurut Sugiyono (2013) penelitian eksperimen merupakan metode yang cocok saat penelitian membutuhkan fokus terhadap variabel terpilih dengan mengontrol variabel lain yang dapat memengaruhi proses. Karena, penelitian ini mengukur persepsi audiens setelah terpapar berita oleh *anchor* AI, eksperimen menjadi pilihan yang cocok karena dibutuhkan kontrol untuk menghindari hal seperti audiens tidak menonton penuh paparan. Riset terdahulu seperti Kim et al. (2022) dan Xue et al. (2022) pun melakukan metode eksperimen dengan memaparkan audiens terhadap *AI newscaster* dengan mengontrol variabel yang dapat memengaruhi variabel terpilih.

Meskipun demikian kontrol diupayakan, riset terdahulu (Jang et al., 2022 ; Sun et al., 2022 ; Kim et al., 2022 ; Xue et al., 2022) juga menyatakan ragam faktor yang masih dapat memengaruhi persepsi audiens terhadap *news anchor* berbasis AI dan kecerdasan buatan pada umumnya. Field dan Hole (2003) menjelaskan bahwa *quasi experimental design* menjadi desain terbaik ketika peneliti tidak mempunyai kontrol penuh terhadap variabel independen. Field dan Hole (2003) menyatakan hanya pengukuran variabel dependen yang dapat dimanipulasi pada desain *quasi experimental*. Terlebih, pada desain *quasi experimental*, cara partisipan dibagi antara kondisi studi tidak dapat dikontrol (Field & Hole, 2003). Riset kini pun sesuai dengan jenis itu, berhubungan peneliti hanya menggunakan kecerdasan buatan dari *tvOne*. Berbeda dengan riset Zhou et al. (2024), penelitian ini tidak mempunyai kontrol terhadap atribut manusia yang diberikan pada *news anchor* AI.

Selebihnya, peneliti akan menggunakan *quasi experimental design* bentuk *Between-Group* oleh Field dan Hole (2003). Penggunaan desain *Between-Group* akan melalui *Posttest only control design* yang merupakan paling lugas dari bentuk

lain. Jenis ini akan membagi kelompok secara acak yang terbagi antara, yang diberikan *treatment* atau disebut sebagai kelompok eksperimen dan, yang tidak diberikan *treatment* atau disebut sebagai kelompok kontrol (Field & Hole., 2003). Namun, peneliti tidak melakukan manipulasi terhadap variabel independen antropomorfisme berhubungan berbagai *tvOne*. Berdasarkan itu, peneliti ini akan mengukur antropomorfisme pada *anchor* AI berbasis *tvOne* antara kelompok pertama yang berupa *anchor*. Versi Lama dengan karakteristik lebih robotik. Sementara itu, kelompok kedua yakni Versi Baru dengan karakteristik yang lebih menyerupai manusia. Setelah terbukti bahwa kedua versi mempunyai perbedaan penilaian antropomorfisme yang signifikan, tiap kelompok akan disebut sebagai kontrol dan eksperimen.

Selebihnya, penelitian eksperimen akan dilakukan secara daring yang juga sebelumnya dilakukan oleh riset terdahulu (Jang et al., 2022). Pada penelitian ini, eksperimen dilakukan dengan cara memasukkan partisipan ke dalam grup *chat* yang menjadi tempat video dikirim sebagai paparan. Sheng et. al, (2018) juga menyatakan keunggulan dari eksperimen virtual yang tidak terbatas ruang dan waktu. Namun, peneliti juga menggarisbawahi keterbatasan yang mungkin timbul dari eksperimen daring seperti perhatian penuh partisipan terhadap paparan *news anchor* AI yang tidak dapat dikontrol.

3.3 Populasi dan Sampel

Semua jenis riset menggunakan sampel karena tidak mungkin ada riset yang dapat memasuki semua orang melakukan semua aktivitas (Huberman, 1994 dalam Punch, 2005). Punch (2005) pun menambahkan untuk riset kuantitatif, kata *sampling* kerap berarti sampel masyarakat. Hal itu terbagi menjadi dua konsep yaitu populasi yang merupakan total target kelompok dalam dunia yang menjadi subjek riset dan ingin mengucapkan suatu hal. Kemudian, sampel yakni kelompok yang ada pada studi dan penjelasan dari siapa data dikumpulkan. Sampel pun bertujuan untuk membuat kesimpulan atau perkiraan terkait hasil yang akan ditemukan jika data dikumpulkan dari total populasi (Field & Hole, 2003).

Berdasarkan Punch (2005), populasi perlu disesuaikan dengan konteks penelitian. Pada penelitian ini, populasi mencakup seluruh konsumen media melalui internet, berhubungan AI berbasis *tvOne* dipublikasi melalui media sosial YouTube dan Instagram. Untuk menggambarkan populasi, sampel didasarkan dari tingkat penetrasi internet Asosiasi Penyelenggara Jasa Internet Indonesia (APJII, 2024). Dinyatakan 221.563.479 jiwa menggunakan internet di Indonesia. Hasil survei APJII (2024) pun, memperlihatkan penetrasi terbesar berasal dari kedua rentang umur yaitu Milenial dengan 93,17 persen dan Gen Z dengan 87,02 persen. Meski Gen Z tidak menjadi terbesar pada penetrasi, tetapi memimpin sebagai kontributor terbesar pada internet dengan angka 34,40 persen yang baru diikuti Milenial dengan 30,62 persen.

APJII (2024) menyatakan juga bahwa masyarakat pendidikan S1 dan SMA/SMK Paket C, mempunyai penetrasi internet 99,91 persen dan 96,31 persen. Angka itu menandakan bahwa rentang usia dengan pendidikan S1 dan SMA hampir secara universal sudah menggunakan internet. Masyarakat dengan pendidikan SMA/SMK pun juga mempunyai kontribusi tertinggi dari lainnya yakni 31,43 persen.

Selebihnya, data APJII (2024) memaparkan bahwa 78,35 persen sampel Gen Z menyatakan “*Video Online*” sebagai konten hiburan yang paling sering dikunjungi, menandakan bahwa rentang usia itu secara mayoritas sudah terbiasa dengan konten video daring. Tidak hanya itu, Gen Z dibanding rentang umur lain, juga menjadi pengguna terbanyak Instagram dengan 51,90 persen. Instagram pun menjadi salah satu media platform yang digunakan *tvOne* untuk mempublikasi konten AI.

Adapun, YouTube sebagai media sosial yang digunakan *tvOne* untuk AI, tetapi hasil APJII (2024) memaparkan bahwa Gen Z tidak secara masih menggunakan YouTube dengan angka 38,63 persen yang berada di bawah Milenial dengan angka 53,42 persen, Baby Boomers dengan 61,63 persen, dan Gen X dengan 62,91 persen. Namun, 93,63 persen seluruh partisipan yang mencakup semua rentang umur menyatakan YouTube sebagai video platform yang paling

sering ditonton. Oleh karena itu, keakraban masyarakat Indonesia terhadap YouTube tidak asing, terlepas dari rentang umur. Berdasarkan itu, kriteria peneliti partisipan akan fokus pada rentang umur Gen Z, khususnya yang merupakan mahasiswa. Definisi Gen Z pun akan mengikuti dari definisi rentang APJII (2024) yakni kelahiran 1997 sampai 2012. Rentang kelahiran itu, juga sesuai dengan definisi Gen Z oleh Pew Research Center (2019).

Penggunaan partisipan mahasiswa untuk eksperimen *anchor* AI juga tidak baru pada penelitian ini. Sebelumnya, salah satu riset utama terkait presenter AI oleh Kim et al.(2022) juga menggunakan mahasiswa sebagai partisipan. Namun, tidak bisa dipungkiri bahwa mahasiswa sebagai partisipan dapat menimbulkan keterbatasan yakni tidak menggambarkan konsumen berita secara umum karena faktor seperti umur dan pengetahuan (Jang et al.,2022 ; Sun et al.,2022). Berdasarkan itu, penargetan kuota dan karakteristik yang disesuaikan dari populasi disebut sebagai *quota sampling* (Neuman, 2014)

Di sisi lain, sampel yang terdiri dari mahasiswa Gen Z, dapat berpengaruh dalam durasi video yang akan dipaparkan yang disebabkan rentang perhatian mereka yang menurun (Chen et al., 2015). Gen Z juga mempunyai penggunaan Instagram yang tinggi, tetapi fitur *reels* yang berbentuk seperti video TikTok dapat mengurangi rentang perhatian orang-orang menjadi beberapa menit hingga paling kecil 15 detik (Bradbury, 2016). Chen et al. (2015) juga menemukan bahwa audiens pada risetnya cenderung memilih video yang berdurasi panjang jika berkualitas tinggi. Pada riset itu, kata “video panjang” diberikan untuk durasi yang rata-rata adalah 2 menit 5 detik, sedangkan “video pendek” durasi yang rata-ratanya adalah 24 detik. Penelitian ini pun akan menjadikan durasi paparan selama 1 menit 28 detik berhubungan kedua video paparan yang digunakan mengakhiri bagian bahasan pada stempel waktu itu. Hal itu pun sesuai dengan temuan Chen et al. (2015) yaitu menyoroti preferensi terhadap kualitas yang mendorong untuk menonton lebih lama, tetapi masih disesuaikan dengan karakteristik partisipan.

Setelah kriteria sudah jelas, partisipan yang mengikuti eksperimen akan berjumlah $n \geq 30$ yang didasarkan dari Central Limit Theorem (CLT). Neuman

(2014) menyatakan CLT adalah hubungan matematis yang berarti saat banyak sampel acak ditarik dari populasi, maka distribusi normal akan terbentuk. Selebihnya, Roscoe dalam Sugiyono (2013) menjelaskan bahwa ukuran sampel yang layak dalam penelitian adalah 30 sampai 500. Selebihnya, saat dibagi menjadi kelompok, tiapnya perlu berisi minimal 30 partisipan. Field dan Hole (2003) juga menyatakan semakin besar jumlah sampel, maka potensi untuk mendeteksi efek asli akan lebih besar. Dinyatakan, banyak eksperimen psikologis tidak menggunakan partisipan yang cukup sehingga berpeluang untuk tidak menghasilkan statistik yang signifikan (Field & Hole., 2003). Penelitian ini pun akan mengumpulkan paling sedikit 60 partisipan yang akan dibagi menjadi 30 untuk kelompok terpapar *anchor AI tvOne* Versi Lama dan 30 untuk kelompok terpapar Versi Baru *anchor AI tvOne*. Pembagian menjadi kelompok itu juga perlu diacak untuk memastikan perbedaan yang sistematis antara kelompok hanya berupa manipulasi variabel independen. Hal itu membantu untuk meminimalisir faktor yang dapat memengaruhi eksperimen (Field & Hole., 2003).

Berdasarkan itu, pengacakan pada penelitian ini dimulai dengan cara mengumpulkan nama-nama partisipan melalui formulir daring. Setelah itu, nama akan diacak menggunakan Microsoft Excel menggunakan rumus “=RAND” untuk menghasilkan nomor yang selalu diacak tiap melakukan aksi di Excel. Lebih lanjut, kolom angka acak bersama daftar nama partisipan diurutkan dari kecil ke besar. Daftar nama pun akan mengikuti urutan angka acak. Daftar nama awal sampai tengah kemudian dipindahkan ke kolom di bawah tulisan “Kelompok 1”. Lalu, Kolom dengan daftar nama yang tersisa akan dinamakan “Kelompok 2”. Pada akhirnya, nama-nama partisipan akan dimasukkan ke kelompok mengikuti label daftar itu.

3.4 Operasionalisasi Variabel/Konsep

Pada penelitian ini, variabel pertama adalah Antropomorfisme dari Godspeed Bartneck (2009). Bartneck (2009) menyatakan bahwa indikator antropomorfisme menggunakan lima item untuk mengukurnya yaitu

Gadungan/Natural, Layaknya mesin/Layaknya manusia, Tidak berkesadaran/Berkesadaran, Artifisial/Menyerupai kehidupan, dan Pergerakan kaku/Pergerakan elegan. Pada penelitian ini item itu menjadi indikator mengikuti antropomorfisme sebagai variabel. Lebih lanjut, indikator “Artifisial/Menyerupai” tidak digunakan karena sudah mempunyai penjelasan yang sama sudah dicakup indikator “Layaknya mesin/manusia”. Adapun, indikator “Gadungan (*Fake*)/Natural” yang dirancang untuk mengacu pada suara *anchor* AI merujuk dari Levy-Landesberg dan Cao (2024). Berdasarkan itu, tiap indikator antropomorfisme menjadi distingtif dengan fokus yang dibedakan dari suara, kemiripan dengan robot atau manusia, kesadaran yang dipersepsikan, dan pergerakan.

Seungguhnya, variabel *audience engagement* Zimmerman et al. (2024) mempunyai tiga indikator yaitu *cognitive engagement*, *affective engagement*, dan *behavioral engagement*. Namun, *cognitive engagement* tidak digunakan untuk menghindari latar belakang pengetahuan audiens terkait topik. *Behavioral* juga dilewatkan berhubungan riset Zimmerman et al. (2024) menggunakan dimensi tersebut, untuk mengukur sikap aksi langsung setelah paparan. Berdasarkan itu, dalam upaya untuk menggunakan variabel yang tidak dipengaruhi faktor luar dan menyesuaikan dengan topik penelitian ini, *affective engagement* menjadi satu-satunya indikator yang digunakan dari variabel *audience engagement*.

Variabel terakhir adalah *Intention to Watch*. Berbeda dengan variabel-variabel sebelumnya, *Intention to Watch* berbentuk sebagai indikator sendiri. Oleh sebab itu, *Intention to Watch* yang digunakan oleh Xue et al. (2022), langsung diperluas menjadi empat item tanpa perubahan.

Tabel 3. 1 Operasionalisasi Variabel

Variabel	Indikator	Item
Antropomorfisme (Bartneck 2009)	Suara gadungan (<i>Fake</i>)/natural	Irama atau nada bicara pembawa berita AI ini saat membawakan berita Terasa datar dan kaku/Terasa Alami

		<p>Pengucapan kata-kata oleh pembawa berita AI saat membacakan berita</p> <p>Tidak jelas dan robotik/Jelas dan manusiawi</p>
		<p>Jeda bicara dan kontrol napas pada suara pembawa berita AI</p> <p>Terputus-putus dan tidak natural/Mengalir dan sangat alami</p>
	Layaknya mesin/manusia	<p>Wajah pembawa berita AI terlihat</p> <p>Tidak proporsional/Proporsional dan natural</p> <p>Ekspresi wajah pembawa berita AI saat membawakan berita</p> <p>Layaknya robot (kaku)/Layaknya Manusia</p> <p>Secara umum pembawa berita AI yang saya lihat.</p> <p>Tidak menunjukkan karakter atau sifat tertentu/Menunjukkan karakter dan sifat seperti manusia</p>
	Dipersepsikan tidak mempunyai kesadaran/mempunyai kesadaran	<p>Saat membawakan berita, Pembawa berita AI seolah</p> <p>Tidak memahami isi berita/Memahami isi berita</p> <p>Saat membawakan berita, pembawa berita AI ini seolah</p> <p>Tidak menunjukkan emosi relevan dengan konteks berita/Menunjukkan emosi sesuai konteks berita</p>

		<p>Saat membawakan berita, pembawa berita AI ini memberi kesan</p> <p>Tidak memiliki pikiran independen/Memiliki pikiran independen</p>
	Pergerakan kaku/elegan	<p>Gerakan tubuh pembawa berita AI ini terlihat</p> <p>Kaku dan robotik/Luwes dan alami</p>
		<p>Saat membawakan berita variasi gerakan presenter AI ini</p> <p>Terbatas dan kaku/Variatif dan natural</p>
		<p>Saat membawakan berita gerakan pembawa AI terlihat</p> <p>Tidak relevan dengan bahasan/Relevan dengan bahasan</p>
Variabel Zimmerman et al. (2024)	Indikator	Item
<i>Audience engagement</i>	<i>Affective engagement</i>	<p>Saat menonton paparan pembawa berita AI</p> <p>Saya merasa bosan/Saya merasa antusias</p>
		<p>Saat menonton paparan pembawa berita AI</p> <p>Saya merasa cemas/Saya merasa tenang</p>
		<p>Saat menonton paparan pembawa berita AI</p> <p>Saya merasa frustrasi/Saya menikmati seluruh paparan</p>
Variabel	Indikator	Item

Xue et al. (2022)		
<i>Intention to watch</i>		Dari skala 1 - 5 seberapa ingin Anda menyaksikan pembawa berita AI ini lagi?
		Tidak ingin/Sangat ingin
		Seberapa setuju Anda dengan pernyataan ini: "Saya akan aktif mencari berita yang dibawakan pembawa berita AI ini"
		Seberapa setuju Anda dengan pernyataan ini? "Saya berminat untuk jadikan pembawa berita AI ini sebagai salah satu sumber berita saya"
		Saya akan berusaha untuk tidak melewatkan berita yang dibawa pembawa berita AI ini

Sumber: Olahan Penulis (2025)

3.5 Teknik Pengumpulan Data

Sebelum data dikumpulkan, peneliti melakukan penyebaran poster melalui media sosial Instagram, X, dan Facebook serta aplikasi *instant messaging* seperti LINE dan WhatsApp. Pada poster dicantumkan tautan dan kode QR registrasi yang mengandung pernyataan persetujuan untuk menjadi partisipan eksperimen. Selibuhnya, tiap orang yang registrasi juga diwajibkan untuk memasuki grup WhatsApp utama. Partisipan yang sudah memasuki grup utama, akan dibagi menjadi dua kelompok secara acak menggunakan Microsoft Excel. Tiap anggota akan secara acak mendapatkan antara "Group 1" atau "Group 2". Partisipan pada grup besar pun diwajibkan untuk memasuki pembagian grup tersebut. Setelah itu, tiap grup akan dikirimkan paparan dan pemberitahuan untuk menonton sampai selesai sebelum mengisi kuesioner.

Pada desain *posttest only control design*, kuesioner akan diberikan setelah eksperimen. Werang (2020) menjelaskan kuesioner sebagai teknik pengumpulan data yang bersumber dari orang yang menjawab pertanyaan penelitian. Hal itu tepat

digunakan dalam konteks penelitian bersampel besar karena mempermudah penelitian dengan mempunyai pertanyaan yang sama dengan tiap partisipan.

Pada penelitian ini, jenis kuesioner menggunakan dua ukuran yaitu Skala Diferensial Semantik menyesuaikan dengan indikator Bartneck (2009) dan Skala Likert. Osgood dalam Hahn and Heit (2015) mendefinisikan Skala Diferensial Semantik sebagai pengukuran yang menggunakan teknik psikometri untuk menghitung rentang psikologis terhadap konsep. Diferensial Semantik pun diatur sebagai garis tunggal antara dua kata sifat yang berlawanan dengan tiap ujung merepresentasikan sifat “sangat negatif” dan “sangat positif”.

Adapun, penggunaan Skala Likert untuk indikator *Intention to Watch*. Sugiyono (2013) pun menyampaikan bahwa Skala Likert tepat digunakan dalam mengukur sikap, pendapat, dan persepsi seseorang atau kelompok terhadap fenomena sosial. Berhubungan, tingkat *affective engagement* dan intensi merupakan persepsi dari partisipan, skala Likert digunakan. Berikut merupakan model Skala Likert pada Sugiyono (2013) berdasarkan definisi Werang (2020).

Tabel 3. 2 Skala Likert

Poin	Keterangan
5	Sangat Setuju (SS)
4	Setuju (S)
3	Ragu-ragu (RG)
2	Tidak Setuju (TS)
1	Sangat Tidak Setuju (STS)

3.6 Teknik Pengukuran Data

Memasuki tahap analisis data, tiap indikator yang tertera akan disingkat untuk memudahkan peneliti dan pembaca. Karena, antropomorfisme (Bartneck,

2009) diukur dengan empat indikator, tiapnya akan ditulis sebagai “Antrop” diikuti dengan abjad dan nomor item. Selbihnya, *affective engagement* merupakan satu dari tiga indikator *audience engagement* (Zimmerman et al., 2024) sehingga akan direpresentasikan sebagai “Engagement A”. Terakhir, *intention to watch* oleh Xue et al., (2022) terlihat terlalu panjang dalam analisis data sehingga disingkat menjadi “IW” diikuti nomor item.

3.6.1 Uji Validitas

Arikunto dalam Werang (2020) menyebutkan validitas sebagai bentuk pengukuran yang menunjukkan tingkat keandalan suatu alat ukur. Uji validitas instrumen pun dapat dilakukan dengan uji validitas isi (*content validity*). Validitas isi dilihat dari tingkat kepastian makna item-item pada kuesioner sudah cukup merepresentasikan konsep (Sekaran dalam Werang, 2020). Menggunakan perangkat JAMOVI, uji *Correlation Matrix* akan digunakan. Merujuk dari Field dan Hole (2003) validitas dapat dievaluasi dengan mengukur korelasi antara ukuran lain dalam konstruk yang sama. Untuk melakukan uji tersebut, penelitian mengumpulkan 30 orang untuk mengikuti uji validitas dan reliabilitas. Berdasarkan, Werang (2020) nilai korelasi pun akan mengikuti tingkatan berikut.

0,800 – 1,000 : Sangat tinggi.

0,600 – 0,799 : Tinggi.

0,400 – 0,599 : Cukup tinggi.

0,200 – 0,399 : Rendah.

0,000 – 0,199 : Sangat rendah

Adapun, George dan Mallery (2021) yang menyebutkan angka paling kecil untuk diterima saat menguji validitas adalah $r = 0,5$. Namun, sebelum mengukur *Correlation Matrix*, penting untuk melakukan uji normalitas untuk mengetahui langkah selanjutnya.

Tabel 3. 3 Uji Normalitas untuk Validitas

Descriptives

	N	Mean	Median	SD	Shapiro-Wilk	
					W	p
Antrop A1	30	2.47	2.00	1.19	0.87	0.002
Antrop A2	30	3.20	3.00	1.12	0.90	0.011
Antrop A3	30	2.87	3.00	1.07	0.84	< .001
Antrop B1	30	3.27	3.00	1.23	0.91	0.015
Antrop B2	30	2.40	2.00	1.07	0.85	< .001
Antrop B3	30	2.43	2.00	1.19	0.86	0.001
Antrop C1	30	3.87	4.00	0.90	0.86	0.002
Antrop C2	30	2.50	2.00	1.19	0.87	0.002
Antrop C3	30	2.30	2.00	0.98	0.86	0.001
Antrop D1	30	2.27	2.00	1.14	0.87	0.002
Antrop D2	30	2.07	2.00	1.08	0.82	< .001
Antrop D3	30	2.90	3.00	1.21	0.90	0.012

Engagement A1	3 0	2.37	2.00	0.99 9	0.87 3	0.002
Engagement A2	3 0	2.93	3.00	1.08 1	0.89 8	0.007
Engagement A3	3 0	3.03	3.00	0.92 8	0.89 3	0.006
IW 1	3 0	2.30	2.00	1.11 9	0.88 4	0.004
IW 2	3 0	1.80	2.00	1.06 4	0.70 2	< .00 1
IW 3	3 0	1.83	1.50	1.08 5	0.76 4	< .00 1
IW 4	3 0	1.93	1.50	1.20 2	0.76 9	< .00 1

Untuk menyatakan distribusi data normal *p-value* perlu lebih dari 0,05 (Field & Hole., 2003). Pada *Correlation Matrix*, setidaknya dua variabel perlu bernilai normal untuk menggunakan Pearson's r. Namun, berhubungan tabel 3.3 memperlihatkan semua variabel tidak normal, Spearman's rho lebih lazim untuk digunakan (Field & Hole., 2003).

3.6.1.1 Uji Validitas Antropomorfisme

Tabel 3. 4 Uji Validitas Suara Gadungan (*Fake*)/Natural (Antrop A)

Correlation Matrix

		Antrop A1	Antrop A2	Antrop A3
Antrop A1	Spearman's rho	—		

	p-value	–		
Antrop A2	Spearman's rho	0.586 ***	–	
	p-value	< .001	–	
Antrop A3	Spearman's rho	0.345	0.417 *	–
	p-value	0.062	0.022	–

Note. * p < .05, ** p < .01, *** p < .001

Pada indikator “Suara gadungan(*Fake*)/natural”

Tabel 3. 5 Uji Validitas Layaknya Mesin/Manusia (Antrop B)

Correlation Matrix

		Antrop B1	Antrop B2	Antrop B3
Antrop B1	Spearman's rho	–		
	p-value	–		
Antrop B2	Spearman's rho	0.526 **	–	
	p-value	0.003	–	
Antrop B3	Spearman's rho	0.321	0.367 *	–

p-value 0.084 0.046 -

Note. * p < .05, ** p < .01, *** p < .001

Tabel 3. 6 Uji Validitas Dipersepsikan Berkesadaran/Tidak berkesadaran (Antrop C)

Correlation Matrix

		Antrop C1	Antrop C2	Antrop C3
Antrop C1	Spearman's rho	-		
	df	-		
	p-value	-		
Antrop C2	Spearman's rho	0.287	-	
	df	28	-	
	p-value	0.124	-	
Antrop C3	Spearman's rho	0.020	0.207	-
	df	28	28	-
	p-value	0.918	0.273	-

Note. * p < .05, ** p < .01, *** p < .001

Tabel 3. 7 Uji Validitas Pergerakan Kaku/Elegan (Antrop D)

Correlation Matrix

		Antrop D1	Antrop D2	Antrop D3
Antrop D1	Spearman's rho	—		
	df	—		
	p-value	—		
Antrop D2	Spearman's rho	0.805 ***	—	
	df	28	—	
	p-value	< .001	—	
Antrop D3	Spearman's rho	0.521 **	0.447 *	—
	df	28	28	—
	p-value	0.003	0.013	—

Note. * p < .05, ** p < .01, *** p < .001

3.6.1.2 Uji Validitas Audience Engagement

Tabel 3. 8 Uji Validitas *Affective Engagement* (Engagement A)

Correlation Matrix

		Engagement A1	Engagement A2	Engagement A3
Engagement A1	Spearman's rho	—		
	df	—		
	p-value	—		

Engagement A2	Spearman's rho	0.528 **	—	
	df	28	—	
	p-value	0.003	—	
Engagement A3	Spearman's rho	0.491 **	0.667 ***	—
	df	28	28	—
	p-value	0.006	< .001	—

Note. * $p < .05$, ** $p < .01$, *** $p < .001$



3.6.1.3 Uji Validitas Intention to Watch

Tabel 3. 9 Uji Validitas *Intention to Watch* (IW)

Correlation Matrix

		IW 1	IW 2	IW 3	IW 4
IW 1	Spearman's rho	—			
	df	—			
	p-value	—			
IW 2	Spearman's rho	0.654 ** *	—		
	df	28	—		
	p-value	< .001	—		
IW 3	Spearman's rho	0.562 **	0.588 ** *	—	
	df	28	28	—	
	p-value	0.001	< .001	—	
IW 4	Spearman's rho	0.498 **	0.748 ** *	0.708 ** *	—
	df	28	28	28	—
	p-value	0.005	< .001	< .001	—

Note. * p < .05, ** p < .01, *** p < .001

Berdasarkan paparan *Correlation Matrix*, Antrop A3, B3, dan keseluruhan Antrop C tidak mencapai Spearman's rho = 0,5 yang berarti tidak valid. Oleh sebab itu, Antrop A3, B3, dan keseluruhan Antrop C tidak digunakan pada eksperimen asli.

3.6.2 Uji Reliabilitas

Hasan dan Malhotra dalam Werang (2020) menyatakan reliabilitas sebagai tingkat konsistensi hasil suatu pengukuran. Semakin tinggi reliabilitas, semakin tinggi juga peluang alat ukur mendapatkan hasil serupa ketika digunakan di penelitian lain. Berdasarkan itu, uji reliabilitas penting untuk mengetahui potensi konsistensi alat ukur dalam penelitian ini. Umumnya, batas minimal indikator reliabel adalah Cronbach's Alpha = 0,70. Namun dalam penelitian eksplanatori dapat turun menjadi 0,60 (Hair dalam Werang 2020). Adapun tingkatan nilai Cronbach Alpha saat menguji reliabilitas (George & Mallery, 2021) sebagai berikut.

Alpha > .9—*excellent*

Alpha > .8—*good*

Alpha > .7—*acceptable*

Alpha > .6—*questionable*

Alpha > .5—*poor*

Alpha < .5—*unacceptable*

Namun, tidak satupun interpretasi nilai Cronbach's Alpha yang diterima secara universal. Nilai alpha pun kerap mengalami inflasi karena terlalu banyak jumlah item yang digunakan. Sebaliknya, ketika item yang digunakan berjumlah sedikit nilai Alpha akan deflasi (George & Mallery, 2021) Berhubungan penelitian ini menggunakan hanya tiga item yang kecil untuk tiap indikator, reliabilitas yang berjumlah > 0.65 masih akan diterima berhubungan sudah mendekati "acceptable" 0.7 dan belum di bawah < 0.6. Uji reliabilitas pun penting untuk mengetahui potensi konsistensi alat ukur dalam penelitian ini.

3.6.2.1 Uji Realibilitas Antropomorfisme

Tabel 3. 10 Uji Reliabilitas Suara Gadungan (*Fake*)/Natural

Scale Reliability Statistics

	Cronbach's α	McDonald's ω
scale	0.729	0.743

Item Reliability Statistics

	Mean	Item-rest correlation	If item dropped	
			Cronbach's α	McDonald's ω
Antrop A1	2.47	0.592	0.593	0.593
Antrop A2	3.20	0.634	0.540	0.543
Antrop A3	2.87	0.440	0.764	0.765



Tabel 3. 11 Uji Reliabilitas Layaknya Mesin/Manusia

Scale Reliability Statistics

	Cronbach's α	McDonald's ω
scale	0.657	0.673

Item Reliability Statistics

	Item-rest correlation	If item dropped	
		Cronbach's α	McDonald's ω
Antrop B1	0.523	0.481	0.483
Antrop B2	0.513	0.509	0.509
Antrop B3	0.379	0.679	0.683



Tabel 3. 12 Dipersepsikan Tidak Sadar/Sadar

Scale Reliability Statistics

	Cronbach's α	McDonald's ω
scale	0.445	0.530

Item Reliability Statistics

	Item-rest correlation	If item dropped	
		Cronbach's α	McDonald's ω
Antrop C1	0.273	0.354	0.359
Antrop C2	0.362	0.157	0.157
Antrop C3	0.198	0.471	0.485



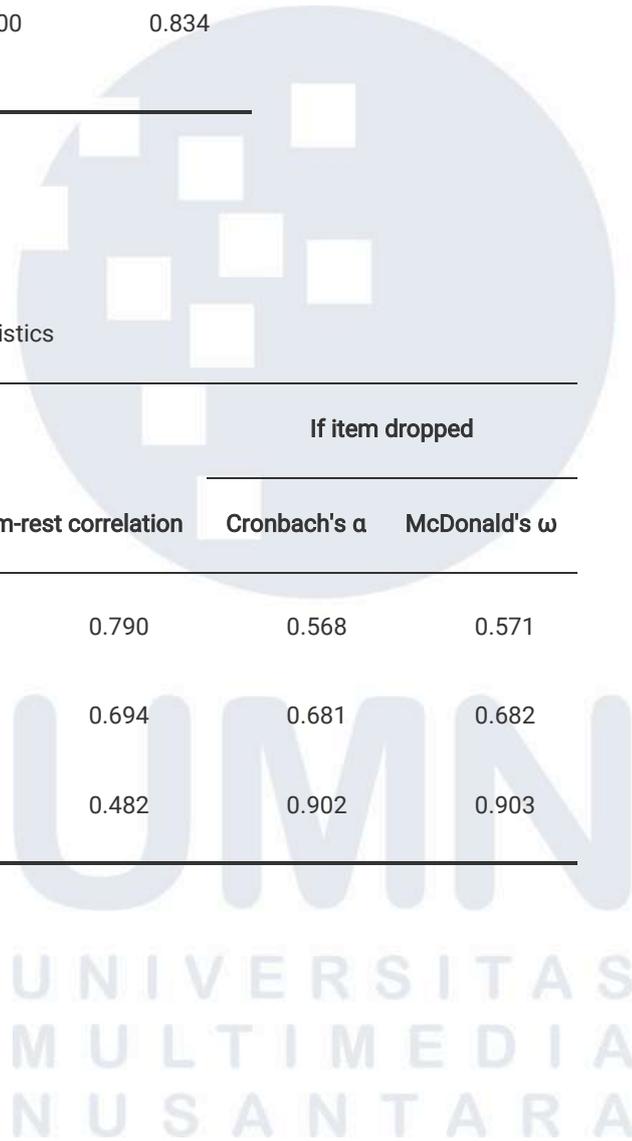
Tabel 3. 13 Uji Reliabilitas Pergerakan Kaku/Elegan

Scale Reliability Statistics

	Cronbach's α	McDonald's ω
scale	0.800	0.834

Item Reliability Statistics

	Item-rest correlation	If item dropped	
		Cronbach's α	McDonald's ω
Antrop D1	0.790	0.568	0.571
Antrop D2	0.694	0.681	0.682
Antrop D3	0.482	0.902	0.903



3.6.2.2 Uji Realibilitas Audience Engagement

Tabel 3. 14 Uji Reliabilitas *Affective Engagement*

Scale Reliability Statistics

	Cronbach's α	McDonald's ω
scale	0.787	0.798

Item Reliability Statistics

	Item-rest correlation	If item dropped	
		Cronbach's α	McDonald's ω
Engagement A1	0.530	0.811	0.817
Engagement A2	0.691	0.638	0.639
Engagement A3	0.674	0.667	0.669

3.6.2.3 Uji Realibilitas Intention to Watch

Tabel 3. 15 Uji Reliabilitas *Intention to Watch*

Scale Reliability Statistics

	Cronbach's α	McDonald's ω
scale	0.914	0.917

Item Reliability Statistics

	Item-rest correlation	If item dropped	
		Cronbach's α	McDonald's ω
IW 1	0.711	0.920	0.921
IW 2	0.872	0.866	0.872
IW 3	0.826	0.881	0.891
IW 4	0.817	0.885	0.887

Jika ada item dengan Cronbach's Alpha yang tidak mencapai 0,3 (Hair dalam Werang 2020), George dan Mallery (2021) anjurkan untuk dihapus. Penghapusan pun dapat meningkatkan reliabilitas keseluruhan indikator. Namun, ada kemungkinan indikator tidak dapat ditingkatkan walaupun item dihilangkan. Pada penelitian ini, Indikator "Antrop C" telah gagal mencapai reliabilitas minimal 0,60 (Hair dalam Werang 2020) sehingga perlu dihilangkan.

3.7 Teknik Analisis Data

Setelah membersihkan dataset yang didapatkan dari kelompok versi lama dan kelompok versi baru. Tiap indikator berdasarkan hasil partisipan perlu dijadikan indeks yaitu rata-rata keseluruhan indikator. Hal itu pun menunjukkan perbedaan deskriptif tiap variabel antara dua kelompok yang diuji. Setelah itu, uji inferensial dengan melakukan uji normalitas terlebih dahulu sebelum memasuki tahap analisis selanjutnya (Field & Hole, 2003). Sama dengan yang dilakukan pada uji validitas, normalitas dilihat dari Shapiro's *p-value* yang sama dengan atau melebihi 0,05 (Field & Hole, 2003). Jika normalitas mencapai *p-value* minimal 0,05, tahap analisis akan dilakukan secara parametrik. Sebaliknya, jika *p-value* kurang dari 0,05 uji analisis akan dilakukan secara non-parametrik. Bersama dengan normalitas, uji homogenitas juga dilakukan sebagai langkah pengujian statistik lanjutan untuk mengetahui tahap analisis akan dilakukan secara parametrik atau non-parametrik. Field dan Hole (2003) pun menyatakan uji parametrik dilakukan saat data memiliki variansi homogen yang berarti dapat dipastikan tidak memiliki variasi yang jauh beda. Sama seperti normalitas, data dinyatakan homogen saat sama atau melebihi *p-value* 0,05..

Setelah normalitas dan homogenitas sudah menetapkan sifat parametrik analisis, *Independent Sample T-test* atau *Mann-Whitney U* dapat dilakukan untuk mencari tahu signifikansi perbedaan dua kelompok terhadap variabel (Field dan Hole., 2003). Namun, penggunaan *Independent Sample T-test* ataupun *Mann-Whitney-U* tidak sesuai pada penelitian ini, karena menggunakan dua variabel yang dipengaruhi. Oleh sebab itu, jika uji *Independent Sample T-test* atau *Mann-Whitney U* digunakan, peneliti perlu melakukannya dua kali untuk tiap variabel. Hal itu pun dapat meningkatkan kemungkinan kesalahan pada hasil data (Field dan Hole, 2003). Pada konteks penelitian ini, pengaruh antropomorfisme dengan variabel dipengaruhi dapat terlihat signifikan dengan angka yang lebih kecil dari 0,05, tetapi ada kemungkinan bahwa hal itu tidak memperlihatkan angka yang akurat. Berdasarkan itu, setelah *Independent Sample T-test* atau *Mann-Whitney U*, analisis inferensial uji beda perlu dilakukan lebih lanjut.

Correlation matrix pun dapat dilakukan untuk tiap indeks variabel untuk mengetahui signifikansi korelasi atau hubungan antara variabel. Namun, perlu digaris bawahi bahwa *correlation matrix* hanya memaparkan hubungan dan belum mencapai pengaruh.

Field dan Hole (2003) menyatakan *Multivariate Analysis of Variance* (MANOVA) lebih cocok digunakan untuk penelitian dengan dua variabel dependen. Pada penelitian ini, uji MANOVA dilakukan mengingat variabel independen dites terhadap dua kelompok dengan kondisi yang berbeda. George dan Mallery (2021) pun menyatakan saat menggunakan uji MANOVA, *p-value* 0,01 atau lebih kecil diharapkan karena *p-value* merupakan probabilitas hubungan yang terjadi tidak ada. Jika $p = 0,01$ berarti hanya ada 1 dari 100 kemungkinan hubungan antara variabel hanya terjadi karena kemungkinan. Jika *p-value* dari uji MANOVA adalah 0,01 atau lebih kecil, dapat diyakinkan bahwa Hipotesis Null (H_0) dapat ditolak. Namun, George dan Mallery (2021) menyampaikan umumnya *p-value* di bawah 0,05 sudah dapat diterima sebagai signifikan. *P-value* yang dilihat pada MANOVA adalah Pillai's Trace karena dianggap sebagai uji terbaik secara kekuatan dan kokohnya statistik (George & Mallery, 2021). Artinya, ada sugesti bahwa nilai *p* Pillai's Trace reliabel walaupun tidak lolos uji asumsi seperti homogenitas.

Analisis terakhir adalah uji mediasi simpel menggunakan *plugin* tambahan JAMOMI bernama *medmod* dari Ravi Selker. Mediasi umumnya dilakukan dengan membandingkan data yang korelasional. Namun, semua bentuk analisis kekuatan mediasi, seutuhnya tidak menjelaskan efek kausal (Hayes, 2005.) Untuk memperkuat hasil mediasi, *Bootstrapping* yang merupakan metode komputasi intensif dapat dilakukan. Dengan melakukan *Bootstrapping*, uji mediasi tidak menghasilkan asumsi distribusi *sampling*, yang umumnya terjadi pada sampel kecil (Hayes, 2005). Pada perangkat JAMOMI pun, peneliti akan melakukan sebesar 1000 sampel saat menggunakan *Bootstrapping*.