

PENERAPAN VERIFIKASI DALAM PROSES PENGEMBANGAN ALAT UKUR PSIKOLOGI

Haryanti Mustika

Fakultas Psikologi, Universitas Gadjah Mada
haryantimustika@mail.ugm.ac.id

Abstrak

Verifikasi artinya pemeriksaan tentang suatu kebenaran atas laporan, pernyataan, dan lain-lain. Verifikasi dalam pengembangan ini tetap memiliki makna yang sama namun pemeriksaan dalam konteks pengembangan alat ukur ini lebih spesifiknya yaitu melakukan pengujian pada alat ukur yang dikembangkan sehingga alat ukur psikologi tersebut dapat dipercaya keabsahannya. Menurut penjelasan tersebut maka dalam tahap penyusunan skala psikologi, verifikasi ini berada dalam tahapan analisis item. Karena analisis item merupakan proses pengujian parameter-parameter item guna mengetahui apakah item memenuhi persyaratan psikometris untuk disertakan sebagai bagian dari skala. Dalam prosedur konstruksi atau penyusunan tes, diawali dengan pengujian karakteristik masing-masing item (daya diskriminasi item, kesukaran item, dan efektivitas distraktor) yang akan menjadi bagian tes yang bersangkutan kemudian diakhiri pengujian estimasi terhadap validitas dan reliabilitas. Secara singkat, pengujian-pengujian tersebut atau analisis item yang dilakukan tersebut adalah salah satu implikasi dalam penerapan pandangan filsafat positivisme yaitu verifikasi dalam pengembangan alat ukur psikologi.

Kata Kunci: verifikasi; alat ukur psikologi

Abstract

Verification means checking the truth of reports, statements, and others. Verification in this development still has the same meaning, but the examination in the context of the development of this measuring instrument is more specific, namely testing the measuring instrument developed so that the psychological measuring instrument can be trusted for its validity. According to this explanation, in the stage of preparing the psychological scale, this verification is in the item analysis stage. Because item analysis is the process of testing item parameters to find out whether the item meets the psychometric requirements to be included as part of the scale. In the construction procedure or test preparation, it begins with testing the characteristics of each item (item discrimination power, item difficulty, and distractor effectiveness) which will be part of the test concerned and then ends with testing estimates of validity and reliability. In short, the tests or item analysis carried out is one of the implications in the application of the positivist philosophical view, namely verification in the development of psychological measuring tools.

Keywords : verification; psychological measurement tool

PENDAHULUAN

Filsafat ilmu memiliki banyak pandangan terhadap ilmu pengetahuan, bidang filsafat ilmu juga membahas terkait kerangka dasar teori keilmuan. Kerangka dasar teori keilmuan dalam filsafat ilmu diketahui ada beberapa tokoh beserta pandangannya masing-masing terhadap keilmuan itu sendiri. Salah satunya adalah kelompok lingkaran wina dengan tokoh filsufnya yang cukup terkenal yaitu Rudolf Carnap (1891-1970). Rudolf Carnap ialah tokoh paling terkenal pada jamannya dan hasil dari pemikirannya yaitu metode verifikasi yang banyak menarik perhatian.

Verifikasi (Muslih, 2016) berasal dari bahasa inggris, yakni *Verification* yang artinya pemeriksaan tentang suatu kebenaran atas laporan, pernyataan, dan lain-lain. Verifikasi merupakan salah satu cara pengujian hipotesis yang tujuan utamanya adalah untuk menemukan teori-teori, prinsip-prinsip, generalisasi, dan hukum-hukum. Menurut Moritz Slick, verifikasi

merupakan pengamatan empiris secara langsung, artinya pernyataan yang diambil langsung dari objek yang diamati itulah yang benar-benar mengandung makna sehingga pengetahuan dimulai dari suatu pengamatan peristiwa.

Prinsip-prinsip verifikasi dalam lingkaran wina yaitu, pertama, mencari garis pemisah antara pernyataan yang bermakna (*meaningfull*) dan pernyataan yang tidak bermakna (*meaningless*) berdasarkan kemungkinan untuk diverifikasi. Kedua, prinsip verifikasi menyatakan bahwa suatu proposisi adalah bermakna jika ia dapat diuji dengan pengalaman dan dapat diverifikasi dengan pengamatan (observasi) (Muslih, 2016). Verifikasi Carnap ini termasuk dalam aliran positivisme logis yaitu sebuah paradigma filsafat ilmu pengetahuan yang berasal dari Lingkaran Wina pada tahun 1920-an. Positivisme logis menurut lingkaran wina yaitu ilmu pengetahuan harus memberikan kriteria yang ketat untuk menetapkan apakah sebuah pernyataan adalah benar atau salah.

Lahirnya paham positivisme banyak masuk ke dalam ranah ilmu yaitu ranah ilmu sosial, ranah ilmu hukum, dan yang utama adalah ranah ilmu eksakta karena filsafat positivisme bersandar kepada suatu hal yang bersifat real, nyata dan kasat mata serta tidak mengacu dari hal yang bersifat metafisik (Herlambang, 2019). Sama halnya dengan pengukuran psikologi, pengukuran psikologi adalah pengukuran aspek-aspek tingkah laku yang nampak, yang dianggap dapat mencerminkan prestasi, bakat, sikap dan aspek-aspek kepribadian yang lain (Muhid, Suhadiyanto, & Nurhidayat, 2015). Secara operasional, pengukuran merupakan suatu prosedur perbandingan antara atribut yang hendak diukur dengan alat ukurnya.

Pengukuran merupakan proses kuantifikasi suatu atribut. Pengukuran yang diharapkan akan menghasilkan data yang valid harus dilakukan secara sistematis. Berbagai alat ukur telah berhasil diciptakan untuk melakukan pengukuran atribut dalam bidang fisik seperti berat badan, luas bidang datar, dsb. Namun pengukuran dalam bidang non-fisik, khususnya dalam bidang psikologi masih dalam perkembangan yang mungkin belum pernah mencapai kesempurnaannya (Muhid, Suhadiyanto, & Nurhidayat, 2015). Beberapa tes dan skala psikologi standar yang telah terstandarkan kualitasnya belum dapat dikatakan optimal, oleh sebab itu sampai saat ini masih banyak yang meningkatkan usaha melakukan pengembangan alat ukur psikologi guna mencapai keberhasilan dalam penyusunan dan pengembangan alat-alat ukur psikologi yang lebih berkualitas.

Pengukuran psikologi memiliki berbagai macam jenis alat ukur psikologi untuk digunakan, sampai saat ini tidak terhitung ada berapa banyak alat ukur psikologi yang ada di dunia. Namun dari alat ukur psikologi yang sudah ada, tidak sedikit pula yang mengembangkan alat ukur tersebut menjadi versi lebih baru, lebih baik dan lebih efektif dari penyusunan awal. Untuk mendapatkan pengembangan alat ukur yang lebih bagus dari versi sebelumnya, perlu adanya pengujian sebagai metode analisis agar hasil dari pengembangan alat ukur psikologi tersebut dapat dipercaya penggunaannya.

Pengujian analisis dalam pengembangan alat ukur ini apabila menggunakan paham filsafat ilmu dengan kerangka teori keilmuannya dapat dikatakan juga sebagai verifikasi. Verifikasi dalam pengembangan ini tetap memiliki makna yang sama yaitu pemeriksaan tentang suatu kebenaran atas laporan, pernyataan, dan lain-lain. Namun pemeriksaan dalam konteks pengembangan alat ukur ini lebih spesifiknya yaitu melakukan pengujian pada alat ukur yang dikembangkan sehingga alat ukur psikologi tersebut dapat dipercaya keabsahannya.

Verifikasi dalam pengembangan alat ukur psikologi ini sangat penting dilakukan karena apabila verifikasi sebagai pengujian alat ukur ini tidak dilakukan maka akan ada dampak yang terjadi. Salah satu dampak besar yang akan ditimbulkan dari tidak adanya verifikasi dalam pengembangan alat ukur psikologi ini yaitu alat ukur yang menjadi tidak akurat dalam

melakukan pengukuran suatu tujuan ukur. Seperti misalnya ketika mengukur kemampuan seseorang untuk melompat jauh, bila alat ukurnya tidak dilakukan verifikasi terlebih dahulu sebelum dilakukan pada pengukuran yang sesungguhnya, maka pengukuran kemampuan melompat jauh tersebut tidak akan akurat hasilnya.

Variasi data yang ada misalnya subjek yang melakukan lompatan dengan terpaksa, subjek yang melompat ketika sedang lelah, atau subjek yang sekedar tidak mau berusaha dengan maksimal tentu tidak akan terlihat perbedaannya. Sehingga dengan dilakukannya verifikasi dalam pengembangan alat ukur maka keakuratan pengukurannya akan terlihat dan jelas adanya.

PEMBAHASAN

Verifikasi

Para filsuf pada ‘kelompok’ Lingkaran Wina pada umumnya mencurahkan perhatiannya untuk mencari garis pemisah antara pernyataan yang bermakna (*meaningful*) dan pernyataan yang tidak bermakna (*meaningless*) berdasarkan kemungkinan untuk diverifikasi. Artinya jika suatu pernyataan dapat diverifikasi, maka ia berarti bermakna, sebaliknya jika tidak dapat diverifikasi berarti tidak bermakna. Prinsip verifikasi ini menyatakan bahwa suatu proposisi adalah bermakna jika ia dapat diuji dengan pengalaman dan dapat diverifikasi dengan pengamatan (*observasi*). Sebagai akibat dari prinsip ini, filsafat tradisional haruslah ditolak, karena ungkapan-ungkapannya melampaui pengalaman, termasuk dalam teologi dan metafisika pada umumnya.

Menurut Carnap (Muslih, 2016), ilmu (*science*) adalah sebuah sistem pernyataan yang didasarkan pada pengalaman langsung, dan dikontrol oleh verifikasi eksperimental. Verifikasi dalam ilmu bukanlah pernyataan tunggal, tetapi termasuk sistem dan subsistem pernyataan tersebut. Verifikasi didasarkan atas “pernyataan protokol” (*protocol statements*). Istilah dipahami sebagai pernyataan yang termasuk protokol dasar atau catatan langsung dari suatu pengalaman yang langsung pula.

Carnap selanjutnya membedakan antara verifikasi langsung dan tak langsung. Apabila suatu pernyataan yang menunjukkan sebuah persepsi sekarang, seperti “sekarang saya melihat sebuah lapangan merah dengan dasar biru”, maka pernyataan ini dapat diuji secara langsung dengan persepsi kita sekarang. pernyataan tersebut dapat diverifikasi secara langsung dengan penglihatan. Artinya jika tidak dilihatnya, maka ia terbantah. Sementara terhadap verifikasi tidak langsung, Makanya mereka menganggap kaum empirisme sebagai leluhurnya, seperti: David Hume, John S. Mill, dan Ernst Mach (Muslih, 2016).

Carnap (Muslih, 2016) memberikan jalan lewat deduksi dari pernyataan perseptual. Suatu pernyataan yang mengandung makna teoretis, tidak mungkin diverifikasi dengan menghadirkan image sesuatu, tetapi dengan kemungkinan pendeduksian dari pernyataan perseptual tersebut, karena kemungkinan verifikasi. Kita tidak punya gambaran aktual tentang bidang elektromagnetik di bidang gravitasional, namun pernyataan-pernyataan perseptualnya dapat dideksi dari pernyataan-pernyataan tersebut. Untuk lebih memahami prinsip verifikasi ‘ala Carnap, menarik dilihat perbedaannya mengenai dua hukum dalam ilmu alam, yaitu hukum empiris dan hukum teoretis. Hukum empiris adalah hukum-hukum yang dapat dikonfirmasi secara langsung dengan observasi-empiris. Istilah “*observable*” sering digunakan untuk banyak fenomena yang secara langsung dapat diamati. Dengan kata lain hukum empiris adalah hukum tentang yang kelihatan (*observable*). Sementara hukum teoretis adalah hukum abstrak karena ia merupakan hipotesis. Menurut Carnap, para fisikawan sepakat bahwa hukum yang berhubungan dengan tekanan, volume, dan temperatur suatu gas, adalah hukum-hukum empiris. Di lain pihak, perilaku dari molekul-molekul tunggal adalah hukum teoretis.

Hubungan kedua hukum ini bisa digambarkan, bahwa hukum empiris membantu menjelaskan suatu fakta yang diamati dan untuk memprediksi suatu fakta yang belum diamati. Dan dengan cara yang sama, hukum teoretis membantu untuk menjelaskan perumusan hukum teoritis dan memberikan peluang untuk men-derivasi-kan sebuah hukum empiris yang baru.

Pengembangan Alat Ukur Psikologi

Secara operasional, pengukuran merupakan suatu prosedur perbandingan antara atribut yang hendak diukur dengan alat ukurnya. Terdapat tiga karakteristik pengukuran (Muhid, Suhadiyanto, & Nurhidayat, 2015), yaitu pertama, pengukuran merupakan perbandingan antara atribut yang diukur dengan alat ukurnya. Kedua, hasilnya dinyatakan secara kuantitatif, dan terakhir hasilnya bersifat deskriptif. Objek pengukuran psikologi disebut sebagai *psychological attributes* atau *psychological traits*, yaitu ciri yang mewarnai atau melandasi perilaku.

Perilaku sendiri merupakan ungkapan atau ekspresi dari ciri tersebut, yang dapat diobservasi. Kegiatan pengukuran psikologis juga sering disebut sebagai tes. Tes adalah kegiatan mengamati atau mengumpulkan sampel tingkah laku yang dimiliki individu secara sistematis dan terstandar. Disebut “sampel tingkah laku”, karena tes hanya mendapatkan data pada waktu tertentu serta dalam kondisi dan konteks tertentu. Artinya, pada saat tes berlangsung, diharapkan data yang diperoleh merupakan representasi dari tingkah laku yang diukur secara keseluruhan.

Pengukuran merupakan proses kuantifikasi suatu atribut. Pengukuran yang diharapkan akan menghasilkan data yang valid harus dilakukan secara sistematis. Berbagai alat ukur telah berhasil diciptakan untuk melakukan pengukuran atribut dalam bidang fisik seperti berat badan, luas bidang datar, dsb. Namun, pengukuran dalam bidang non-fisik, khususnya dalam bidang psikologi, masih dalam perkembangan mungkin belum pernah mencapai kesempurnaannya. Beberapa tes dan skala psikologi standar dan yang telah terstandarkan kualitasnya belum dapat dikatakan optimal. Sebab masih terus berkembang seiring dengan pesatnya teori pengukuran, hal ini dimungkinkan untuk meningkatkan usaha guna mencapai keberhasilan dalam penyusunan dan pengembangan alat-alat ukur psikologi yang lebih berkualitas.

Untuk mengembangkan alat ukur skala psikologi, sebagaimana menurut Gable (Muhid, Suhadiyanto, & Nurhidayat, 2015), diperlukan beberapa langkah sebagai berikut: (1) mengembangkan definisi konseptual; (2) mengembangkan definisi operasional; (3) memilih teknik pemberian skala; (4) melakukan review justifikasi butir, yang berkaitan dengan teknik pemberian skala yang telah ditetapkan di atas; (5) memilih format respons atau ukuran sampel; (6) penyusunan petunjuk untuk respons; (7) menyiapkan draft instrumen, (8) menyiapkan instrumen akhir; (9) pengumpulan data uji coba awal; (10) analisis data uji coba dengan menggunakan teknik analisis faktor, analisis butir dan reliabilitas; (11) revisi instrumen; (12) melakukan uji coba final; (13) menghasilkan instrumen; (14) melakukan analisis validitas dan reliabilitas tambahan; dan (15) menyiapkan manual instrumen.

Penerapan Verifikasi dalam Pengembangan Alat Ukur

Tahap-tahap penyusunan skala psikologi atau alur kerja dalam penyusunan skala psikologi (Muhid, Suhadiyanto, & Nurhidayat, 2015) yaitu : (1) Penetapan tujuan, (2) Menetapkan kawasan (*domain*) ukur, (3) Menyusun atribut dan indikator perilaku, (4) Menyusun *blue print*, (5) Menuliskan item, (6) Penskalaan dan penentuan skor, (7) Seleksi item, (8) Uji coba, (9) Analisis item, (10) Kompilasi pertama, (11) Kompilasi kedua, dan tahap terakhir (13) Format akhir. Verifikasi dalam pengembangan alat ukur ini tetap memiliki makna yang sama yaitu pemeriksaan tentang suatu kebenaran atas laporan, pernyataan, dan lain-lain. Namun pemeriksaan dalam konteks pengembangan alat ukur ini lebih spesifiknya yaitu

melakukan pengujian pada alat ukur yang dikembangkan sehingga alat ukur psikologi tersebut dapat dipercaya keabsahannya.

Menurut penjelasan tersebut maka dalam tahap pengembangan maupun penyusunan alat ukur psikologi, verifikasi ini berada dalam tahapan analisis item. Karena analisis item merupakan proses pengujian parameter-parameter item guna mengetahui apakah item memenuhi persyaratan psikometris untuk disertakan sebagai bagian dari skala (Muhid, Suhadiyanto, & Nurhidayat, 2015). Parameter item yang perlu diuji salah satunya adalah daya beda atau daya diskriminasi item yang merupakan pengujian sejauh mana kemampuan suatu item untuk membedakan individu yang satu dari yang lainnya berdasarkan atribut yang diukur oleh tes dinyatakan sebagai parameter daya diskriminasi (*discriminating power*) item (Azwar, Konstruksi Tes Kemampuan Kognitif, 2016).

Tokoh lain juga menyatakan bahwa analisis item adalah suatu kegiatan mengevaluasi item-item alat tes. Dari kegiatan ini diharapkan didesain sebuah alat tes dengan jumlah item minimum, namun reliabilitas dan validitas yang maksimum (Kaplan & Saccuzzo, 2016). Menurut Muhid, dkk., (2015) analisis item dapat dilakukan dengan dua metode yaitu secara kualitatif dan secara kuantitatif. Pertama, secara kualitatif yaitu menyangkut keterwakilan tingkah laku domain menjadi item dalam alat tes (konten dan form) *a content validity* (menyangkut expert judgement). Sedangkan metode selanjutnya adalah secara kuantitatif, dibagi menjadi *item difficulty* (kesukaran item) dan *item discriminant* (diskriminasi item).

Item difficulty merupakan persentase (proporsi) orang yang menjawab item dengan benar, sedangkan *item discriminant* adalah perbandingan antara proporsi orang yang menjawab benar dalam kelompok *upper* dengan proporsi orang yang menjawab benar dari kelompok *lower*. Perbedaan proporsi ini disebut sebagai *indec of discriminant* (D). Sedangkan menurut Azwar (2016), analisis item merupakan evaluasi terhadap item dengan dilakukan --paling tidak-- tiga parameter, yaitu sebagai berikut :

a) Daya Beda (Daya Diskriminasi) Item

Secara prinsip, dijelaskan bahwa daya diskriminasi item dicerminkan oleh perbedaan jawaban terhadap item diantara kelompok subjek yang cerdas dan yang tidak cerdas, atau kelompok subjek yang berbakat dan tidak berbakat, atau antara kelompok yang sudah berhasil dan yang belum berhasil dalam belajar. Kemampuan item untuk membedakan subjek, secara kuantitatif, diindikasikan oleh parameter daya diskriminasi item. Parameter tersebut dapat diestimasi melalui komputasi koefisien korelasi item-total, indeks daya diskriminasi item, dan komparasi mean skor tes.

b) Tingkat Kesukaran Item

Dalam teori skor-murni klasik, yang dimaksudkan sebagai taraf kesukaran item adalah parameter yang mendeskripsikan seberapa sulitkan bagi kelompok subjek yang dites untuk memberikan jawaban yang benar terhadap suatu item. Parameter kesukaran item suatu item merupakan deskripsi kesukaran item tersebut bagi seluruh kelompok subjek, bukan bagi masing-masing subjek secara individual.

Taraf kesukaran suatu item bagi subjek adalah berbeda-beda dan berapa sulit atau berapa mudahnya suatu item bagi seorang subjek tidak diketahui. Apa yang diketahui adalah apabila seorang subjek dapat menjawab benar suatu item maka taraf kesukaran item tersebut lebih rendah daripada taraf kemampuannya dalam menjawab. Sebaliknya, apabila subjek salah dalam menjawab suatu item berarti bahwa tingkat kemampuannya lebih rendah daripada taraf kesukatan item yang bersangkutan.

c) Efektivitas Distraktor

Item yang baik tidak saja memiliki daya diskriminasi yang tinggi dan tingkat kesukaran yang sesuai, tapi juga memiliki distraktor-distraktor yang efektif. Fungsi distraktor berlawanan dengan fungsi diskriminatif item. Apabila daya diskriminasi item diperlihatkan oleh lebih besarnya proporsi subjek yang cakap yang dapat menjawab dengan benar dibandingkan proporsi subjek yang tidak mampu, maka fungsi diskriminatif distraktor yang efektif justru diperlihatkan oleh lebih besarnya proporsi subjek yang tidak cakap yang terjebak oleh distraktor dibandingkan proporsi subjek yang mampu.

Idealnya distraktor harus dipilih hanya oleh subjek yang tidak cakap atau tidak mampu, sedangkan subjek yang cakap tidak ada yang memiliki distraktor. Beberapa cara memeriksa efektivitas distraktor yaitu dengan korelasi poin biserial atau korelasi biserial dan dengan menggunakan komparasi *mean* skor tes.

Analisis ketiga parameter tersebut dilakukan berdasarkan data respon subjek terhadap item-item dalam tes. Setelah ketiga parameter tersebut dilakukan untuk prosedur analisis item, selanjutnya dilakukan juga pengujian untuk mencari indeks validitas item dan indeks reliabilitas item (Azwar, 2019). Berbeda dengan koefisien reliabilitas skor tes yang merupakan statistik konsistensi item-item tes dalam menjalankan fungsi ukurnya secara bersama-sama dalam satu perangkat dan lebih berkaitan dengan varians error, maka indeks reliabilitas item ini merupakan parameter item yang diderivasi dari konsistensi satu item dengan keseluruhan item lainnya yang ada dalam tes yang bersangkutan.

Fungsi indeks validitas dan reliabilitas item dalam analisis item (Azwar, 2017) adalah sebagai bagian dari pertimbangan dalam seleksi item yang bertujuan peningkatan validitas dan reliabilitas secara keseluruhan. Penggunaan indeks validitas dan reliabilitas item patut menjadi pertimbangan dalam pemilihan item apabila reliabilitas dan validitas tes secara keseluruhan belum diketahui atau sudah diketahui dan hasilnya belum cukup memuaskan. Apabila reliabilitas tes secara keseluruhan sudah memuaskan, pemilihan item dapat didasarkan pada koefisien korelasi item-total saja.

Menurut Magnussion, Allen, Crocker, dkk., (Azwar, Konstruksi Tes Kemampuan Kognitif, 2016) Indeks validitas dan reliabilitas item menjadi penting dikarenakan rumusnya yang melibatkan varians skor item. Varians skor item memiliki kontribusi dalam menentukan validitas dan reliabilitas tes. Indeks reliabilitas item berkaitan dengan tingginya reliabilitas tes. Item variansnya lebih besar memiliki kontribusi yang lebih besar pula terhadap reliabilitas tes. Karena itu, menurut Azwar (2016) bila harus memilih satu diantara dua item yang koefisien korelasi item-totalnya sama, hendaklah dipilih yang memiliki variasi yang lebih besar. Sebaliknya, apabila tes berisi item yang variansnya homogen (deviasi standar item kurang lebih setara satu sama lain) maka pemilihan item dapat disandarkan pada besarnya koefisien korelasi item-total saja.

Dalam prosedur konstruksi atau penyusunan tes, diawali dengan pengujian karakteristik masing-masing item (daya diskriminasi item, kesukaran item, dan efektivitas distraktor) yang akan menjadi bagian tes yang bersangkutan kemudian diakhiri pengujian estimasi terhadap validitas dan reliabilitas. Secara singkat, pengujian-pengujian tersebut atau analisis item yang dilakukan tersebut adalah salah satu implikasi dalam penerapan pandangan filsafat positivisme yaitu verifikasi dalam pengembangan alat ukur psikologi.

KESIMPULAN

Verifikasi dalam pengembangan ini tetap memiliki makna yang sama yaitu pemeriksaan tentang suatu kebenaran atas laporan, pernyataan, dan lain-lain. Namun pemeriksaan dalam

konteks pengembangan alat ukur ini lebih spesifiknya yaitu melakukan pengujian pada alat ukur yang dikembangkan sehingga alat ukur psikologi tersebut dapat dipercaya keabsahannya. Menurut penjelasan tersebut maka dalam tahap pengembangan maupun penyusunan alat ukur psikologi, verifikasi ini berada dalam tahapan analisis item. Karena analisis item merupakan proses pengujian parameter-parameter item guna mengetahui apakah item memenuhi persyaratan psikometris untuk disertakan sebagai bagian dari skala (Muhid, Suhadiyanto, & Nurhidayat, 2015). Parameter item yang perlu diuji salah satunya adalah daya beda atau daya diskriminasi item yang merupakan pengujian sejauh mana kemampuan suatu item untuk membedakan individu yang satu dari yang lainnya berdasarkan atribut yang diukur oleh tes dinyatakan sebagai parameter daya diskriminasi (*discriminating power*) item (Azwar, Konstruksi Tes Kemampuan Kognitif, 2016).

Dalam prosedur konstruksi atau penyusunan tes, diawali dengan pengujian karakteristik masing-masing item (daya diskriminasi item, kesukaran item, dan efektivitas distraktor) yang akan menjadi bagian tes yang bersangkutan kemudian diakhiri pengujian estimasi terhadap validitas dan reliabilitas. Secara singkat, pengujian-pengujian tersebut atau analisis item yang dilakukan tersebut adalah salah satu implikasi dalam penerapan pandangan filsafat positivisme yaitu verifikasi dalam pengembangan alat ukur psikologi. Verifikasi dalam pengembangan alat ukur psikologi ini sangat penting dilakukan karena apabila verifikasi sebagai pengujian alat ukur ini tidak dilakukan maka akan ada dampak besar yang akan ditimbulkan dari tidak adanya verifikasi dalam pengembangan alat ukur psikologi ini yaitu alat ukur yang menjadi tidak akurat dalam melakukan pengukuran suatu tujuan ukur, sehingga verifikasi ini dilakukan guna mencapai keberhasilan dalam penyusunan dan pengembangan alat-alat ukur psikologi yang lebih berkualitas.

DAFTAR PUSTAKA

- Azwar, S. (2016). *Konstruksi Tes Kemampuan Kognitif*. Yogyakarta: Pustaka Pelajar.
- Azwar, S. (2017). *Reliabilitas dan Validitas Edisi 4*. Yogyakarta: Pustaka Pelajar.
- Azwar, S. (2019). *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Belajar.
- Herlambang, P. H. (2019). Positivisme dan Implikasi Terhadap Ilmu dan Penegakan Hukum. *Indonesian State Law Review, Vol. 2, No. 1*, 103-110.
- Kaplan, R. M., & Saccuzzo, D. P. (2016). *Pengukuran Psikologi: Prinsip, Penerapan, dan Isu Edisi 7*. Jakarta: Salemba Humanika.
- Muhid, A., Suhadiyanto, & Nurhidayat, D. (2015). *Pengembangan Alat Ukur Psikologi*. Surabaya: Universitas Islam Negeri Sunan Ampel Surabaya.
- Muslih, M. (2016). *Filsafat Ilmu: Kajian atas Asumsi Dasar, Paradigma, dan Kerangka Teori Ilmu Pengetahuan*. Yogyakarta: LESFI.