REVIEW

REVIEW OF THE USE AND UTILITY OF
CENTRALISED EVALUATIONS

Final Report

Evaluation of Finland's Development Policy and Cooperation

2023/2

REVIEW

**REVIEW OF THE USE AND UTILITY OF
CENTRALISED EVALUATIONS**

Final Report

Kristiina Mikkola
Päivi Äijälä

kristiina mikkola
C O N S U L T I N G

2023/2

Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF BOXES

Acronyms and Abbreviations

BEAM	Business with Impact -programme
CCO	Cross-Cutting Objective
CSO	Civil Society Organisations
DAC	Development Assistance Committee
DPC	Development Policy Committee
EDF	European Development Fund
EMS	Evaluation Management Services
EUR	Euro
EVA-11	Development Evaluation Unit
EvalNet	The DAC Network on Development Evaluation
FCR	Findings, Conclusions and Recommendations
FGD	Focus Group Discussion
FINCEED	Finnish Centre of Expertise in Education and Development
HRBA	Human Rights Based Approach
KII	Key Informant Interview
MEUR	Million euro
MFA	Ministry for Foreign Affairs
MIT	Massachusetts Institute of Technology
NAO	National Audit Office
ODA	Official Development Assistance
OECD	Organisation for Economic Co-operation and Development
RBM	Results-Based Management
RQ	Review Question
ToC	Theory of Change
ToR	Terms of Reference

Yhteenveto

1. Tarkastelun tarkoitus, tavoitteet ja metodologia

Keskitettyjen evaluointien käyttöä ja hyödyllisyyttä arvioiva tarkastelu on tehty ulkoministeriön (UM) kehitysevaluoinnin yksikön (EVA-11) toimeksiannosta.

Tarkastelun tarkoitus on vahvistaa evaluointien käyttöä tietopohjaisessa päätöksenteossa, oppimisessa ja tilivelvollisuudessa sekä edistää evaluointien hyödyllisyyttä. Siten tarkastelu vaikuttaa myös kehityspolitiikan ja kehitysyhteistyön kehittämiseen. Tarkastelun tavoitteet ovat:

- selvittää, mitkä päätekijät edistävät tai estävät evaluointitulosten käyttöä ja suositusten toimeenpanoa sekä UM:ssä että UM:n suomalaisten sidosryhmien parissa,

- selvittää pääteemat ja/tai UM:lle strategisesti tarkoituksenmukaiset asiat, joihin liittyvien suositusten toimeenpanossa on merkittäviä aukkoja ja

- esittää selkeitä johtopäätöksiä ja suosituksia, jotka auttavat UM:tä edistämään evaluointien käyttöä ja toimeenpanemaan suosituksia.

Tarkastelu perustuu tiimi- ja yhteiskehittämisen menetelmille painottaen tulevaisuutta. Tarkastelutiimi keräsi tietoa yhteensä 36 tapaamisessa fasilitoituina työpajoina, ryhmäkeskusteluina ja yksilö- tai paritapaamisina marras-joulukuussa 2022. Tietoa saatiin 62 henkilöltä UM:stä ja Suomen edustustoista, Kehityspoliittisesta toimikunnasta (KPT), eduskunnasta, valtiontalouden tarkastusvirastosta, suomalaisista kansalaisjärjestöistä, yksityissektorin organisaatioista ja evaluoinnin ammattilaisilta. Evaluointiraportit, johdonvastineet ja niiden takaisin raportoinnit ja evaluoinnin vuosiraportit muodostivat keskeisen asiakirja-aineiston.

Tarkastelu kohdistui keskitettyihin evaluointeihin, jotka EVA-11 on tilannut v. 2015–2022. Tarkastelujakson aikana on tehty yhteensä 23 evaluointia ja kolme tarkastelua. Ne vaihtelevat maaohjelmaevaluoinneista kehitysyhteistyön instrumenttien evaluointeihin, temaattisiin evaluointeihin ja metaevaluointeihin.

2. Vastaukset tarkastelukysymyksiin

Kysymys 1: Mikä on keskitettyjen evaluointien muutosvoima?

Keskitetyillä evaluoinneilla on ollut vain vähäinen vaikutus politiikkakeskusteluun. Vuonna 2021 valmistuneen kehityspolitiikan ylivaalikautisen selonteon sekä vuosina 2018 ja 2022 valmistuneiden kehityspolitiikan tulosraporttien lisäksi ministeriön ihmiset eivät jakaneet muita esimerkkejä evaluointien hyödyntämisestä poliittisessa vuoropuhelussa.

UM:ssä evaluoinneista keskustellaan kehityspolitiikkaa ja kehitysyhteistyötä toteuttavilla osastoilla ja yksiköissä. EVA-11 laatii evaluoinnin vuosiraportin ja se menee laajennettuun johtoryhmään, mutta ryhmässä ei keskustella yksittäisistä evaluoinneista. Sama koskee kehityspoliittista toimikuntaa, eduskuntaa ja ulkoasiainvaliokuntaa. Ministerien kabinetit saavat tietoa valmistuneista evaluoinneista, mutta niistä ei välttämättä keskustella ollenkaan.

Ulko-, turvallisuus-, kauppa- ja kehityspolitiikan johdonmukaisuus on ollut mukana vain muutamassa evaluoinnissa. Kehityspolitiikka on osa ulko- ja turvallisuuspolitiikkaa, joten politiikkajohdonmukaisuutta voisi parantaa sisällyttämällä eri politiikkalohkot osaksi tulevia evaluointeja.

Tarkastelussa löytyi paljon myönteistä näyttöä siitä, että evaluoinnit ovat vaikuttaneet kehitysyhteistyön toimeenpanoon UM:ssä, erityisesti linjauksiin, ohjeisiin ja prosesseihin. Toimeenpanevia yksikköjä ja osastoja ovat palvelleet sellaiset evaluoinnit, joilla on selkeä ja käytännönläheinen teema. Keskusteluissa nousi usein esille evaluoinnin ajoitus. Parhaat esimerkit evaluoinnin myönteisistä vaikutuksista ovat niitä, joissa evaluointi on tukenut päätöksentekoa tai yhteistyön toteutusta. Esimerkiksi maaohjelmiin ja -strategioihin sekä tulosjohtamiseen liittyneet evaluoinnit ovat olleet sellaisia.

Metaevaluointeja on hyödynnetty evaluointikäsikirjan, evaluointimenetelmien ja koulutussisältöjen kehittämisessä. Vuosina 2015–2022 valmistuneet evaluoinnit eivät ole olleet kehityspolitiikan näkökulmasta katsoen muutosvoimaisia.

Kysymys 2: Miten evaluointeja käytetään ja miten hyödyllisiä ne ovat?

Evaluointien käyttö UM:ssä ja muualla vaihtelee paljon. Aktiiviset käyttäjät löytyvät yksiköistä ja osastoilta, mutta mitä ylemmäs päätöksentekohierarkiassa mennään, sitä vähemmän evaluointeja käytetään.

Ministeriössä evaluoinnit ovat vaikuttaneet erilaisten ohjeistojen kehittämiseen, joista esimerkkeinä mm. maaohjelmia, kolmoisneksusta, yhteistyötä kansalaisyhteiskunnan kanssa ja läpileikkaavia teemoja koskevat ohjeistot. Evaluoinnit ovat hyödyttäneet myös lähestymistapojen kehitystä, kuten tulosperustainen lähestymistapa, riskienhallinta ja ihmisoikeusperustaisuus.

Useat suomalaiset kansalaisjärjestöt hyödyntävät evaluointeja oman järjestön kehittämiseen, oppimiseen ja vaikuttamistyöhön. Kehityspoliittisen toimikunnan jäsenet ja sihteeristö tietävät evaluoinneista ja arvostavat niistä saatavaa tietoa, mutta he eivät käytä niitä tällä hetkellä.

Johdonvastineprosessin tarkoitus on varmistaa evaluointitulosten hyödyntäminen UM:ssä. Prosessia on parannettu viime vuosina, mutta tarkastelu paljasti edelleen puutteita. Prosessi ohjaa etupäässä suositusten työstämiseen mutta huomattavasti vähemmän evaluoinnin muun annin, kuten havaintojen ja johtopäätösten hyödyntämiseen.

Johdonvastineen taulukot sisältävät paljon tietoa ja niitä käytetään EVA-11 ohjeiden mukaan. Silti niistä ei aina ilmene, onko suosituksista löydetty yhteinen ymmärrys. Noin kymmenen prosenttia evaluointien suosituksista on täysin hylätty johdonvastineprosessissa, toisinaan ilman selkeää perustelua. Johdonvastineen parantaminen onkin tarpeen: sisältö, aikataulu ja vastuukysymykset on hyvä pohtia uudelleen.

Johdonvastineista tehtyjen seurantaraporttien mukaan valtaenemmistö sovituista toimenpiteistä on toteutettu sovitussa kahden vuoden määräajassa. Tämä kertoo siitä, että ministeriössä sitoudutaan näiden toimenpiteiden toteutumiseen. Kahden vuoden määräaikaa pidetään kuitenkin liian pitkänä, eikä nykyinen taulukkomuotoinen raportti sisällä analyysiä toimenpiteiden vaikutuksista tai oppimisesta. Raporteista puuttuu tieto evaluointitulosten aikaansaamista positiivisista vaikutuksista.

Kysymys 3: Mitkä tekijät edesauttavat tai estävät keskitettyjen evaluointien käyttämistä?

Tarkastelu arvostaa sitä, miten paljon EVA-11 on panostanut evaluointiprosessin kehittämiseen viime vuosina. Kehitys näkyy työkaluissa, ohjeistuksessa ja koulutuksessa, joita on tehty sekä UM:n sisäisille että ulkoisille osallistujille.

EVA-11 identifioi evaluointiaiheita osallistavalla otteella, mutta viime kädessä EVA-11 päättää mitä evaluoidaan. Viime vuosina EVA-11 on tilannut vain laajoja evaluointeja, joiden aikajänne on pitkä, ideasta valmiiseen raporttiin menee n. kaksi vuotta. Nopeat muutaman kuukauden tarkastelut aloitettiin v. 2021, ja ministeriön ihmiset pitävät niitä hyödyllisinä.

EVA-11 on ulkoistanut evaluointien toteutuksen. Kehitysevaluointipalveluja koskeva puitejärjestely (*Evaluation Management Services, EMS*) on varmistanut sen, että evaluoinnit tehdään standardoidusti ja tasalaatuisesti.

Keskusteluissa monet arvostivat avointa viestintää evaluoinnin toteutuksen aikana ja mahdollisuutta osallistua haastattelun lisäksi referenssiryhmään. Myös työpajaa, jossa keskustellaan evaluoinnin alustavista havainnoista, johtopäätöksistä ja suosituksista, pidettiin hyödyllisenä.

Evaluointien toteutuksen aikana omistajuus on EVA-11:llä ja evaluointitiimillä. Omistajuus ja vastuu evaluointisuositusten toimeenpanosta siirtyy osastolle ja/tai yksikölle siinä vaiheessa, kun raportti on valmis ja johdonvastineprosessi käynnistyy. Toisinaan omistajuuden siirtyminen on ollut hankalaa.

Käyttäjät odottavat evaluoinneilta parannuksia raportointiin, tiedon helppoa saatavuutta ja räätälöityjä tietotuotteita esim. ennakkotietoa, muutaman kohdan päähuomioita, yhden sivun tiivistelmiä jne. Raporteissa on paljon hyödyllistä tietoa, joka ei vanhene nopeasti, joten tietoa voisi hyödyntää muutoinkin kuin johdonvastineprosessissa.

Nykymuotoinen johdonvastine toimii hyvin silloin, kun käsiteltävät asiat ovat sellaisella tasolla, johon osastolla/yksiköllä tai kehityspoliittisella ohjausryhmällä on selkeä mandaatti. Johdonvastine ei kuitenkaan toimi silloin, kun suositukset koskevat ministeriön tai edustustojen henkilöstöä. Puolet tarkasteluajanjakson evaluoinneista esittivät suosituksia ministeriön henkilöstöasioihin, henkilöstön määrään tai osaamisen kehittämistarpeisiin, ja monissa muissa evaluointiraporteissa oli näihin liittyviä havaintoja ja johtopäätöksiä. UM:n hallintopalvelut ei toimi evaluointiraportin pohjalta eikä UM:n johtokaan ole asiassa toiminut. Henkilöstöasioiden ratkaisemattomuus luo merkittävän riskin Suomen kehityspolitiikalle ja -yhteistyölle.

Kysymys 4: Miten UM voisi hyödyntää paremmin evaluointitietoa ja oppeja yhtenä tiedolla johtamisen välineenä?

Tarkastelujakson aikana 2015–2022 evaluointeja hyödynnettiin tietopohjaisessa päätöksenteossa osasto- ja yksikkötasolla UM:ssä, mutta evaluointien potentiaalia tukea tiedolla johtamista läpi koko ministeriön, ml. ministeriön poliittinen ja virkamiesjohto ei ole saavutettu.

Evaluointien käyttö oppimiseen on enemmän yksilöiden varassa kuin laajemmin organisaation oppimista. Tarkastelussa tuli esiin monia esimerkkejä siitä, miten yksilöt käyttävät evaluointeja ja oppivat niistä. Oppimista ja hyödyntämistä haastaa informaatiotulva. UM:n konteksti, kuten henkilöstön rotaatio, erityisvirkahenkilöiden vaihtuminen ja riittämätön henkilöstön määrä suhteessa tehtäviin, vaikuttavat myös kielteisesti evaluoinneista oppimiseen ja tiedon jakamiseen.

Evaluointien referenssiryhmiin osallistuminen on näyttäytynyt huikeana tiimioppimisen polkuna. Innostavat ideat ja yhteistyö on saanut aikaan sen, että ihmiset toteuttavat muutoksia jo evaluoinnin aikana. Johdonvastineprosessiin liittyvästä työryhmästä ei sanottu samaa. Toistaiseksi vain joissakin UM:n yksiköissä ja osastoilla evaluoinnit nähdään oppimisena, muu organisaatio ei tätä ajatusta vielä tunne.

Tarkastelu löysi useita ongelmia evaluointitiedon hyödyntämisessä. Useimmat haastatelluista kertoivat työkuormasta ja kiireisistä päivistä, joten työaikaa oppimiseen, reflektointiin ja uuden tiedon käsittelyyn on vähän, mikä estää evaluointitulosten käyttöä mihinkään tarkoitukseen. Evaluointien tuottama tieto ja tietämys ei ylitä yksikkörajoja ministeriössä. Moni nosti esiin myös vaikeuden löytää evaluointeja UM:n tietojärjestelmistä.

Evaluoinneissa on potentiaalia oppimiseen ja tiedon jalostamiseen. EVA-11:n tiimillä on visio keskitetyistä evaluoinneista merkittävänä osana tiedolla johtamisen edistämistä UM:ssä. Tiimillä on syvällinen tuntemus aiemmista ja tulevista evaluoinneista ja heillä on mahdollisuus vaikuttaa päätöksentekijöihin, koska EVA-11 on edustettuna kaikissa tärkeissä ryhmissä. EVA-11:ssä on myös hyödyllistä kokemusta sellaisista kollektiivisista oppimisratkaisuista, joilla voi edistää evaluoinneista oppimista ja tiedon jalostamiseen yhdessä.

3. Johtopäätökset

Tarkastelujakson aikana EVA-11 on pyrkinyt parantamaan evaluointeja ja evaluointiprosessia osallistamalla ihmisiä enemmän, kehittämällä ohjeita ja työkaluja ja tietotuotteita. Tämä kehitystyö on ollut tarpeen, ja tarkastelu arvostaa ja pitää näitä toimia tarkoituksenmukaisena oppimisen ja tiedolla johtamisen edistämiseksi ministeriössä.

Tarkastelu nostaa esiin kehityskohteita. On tärkeä muistaa, että evaluointien käyttö ja hyödyllisyys on vain osittain EVA-11:n vallassa. Evaluointien hyödyllisyys riippuu osastoista ja yksiköistä, joiden on tarkoitus ottaa omistajuus evaluointituloksista ja hyödyntää niitä käytännössä.

Keskitetyt evaluoinnit ovat vaikuttaneet ja muuttaneet kehitysyhteistyön linjauksia, ohjeistuksia ja prosesseja. Vaikutus poliittiseen vuoropuheluun ja sitä myöten kehityspoliittiseen keskusteluun on rajallinen. Keskitetyt evaluoinnit eivät saavuta poliittisia päättäjiä eduskunnassa tai kehityspoliittisessa toimikunnassa (KPT). Toimikunnan jäsenet tunnistavat evaluointien arvon ja laadun, mutta niitä ei käytetä. Ministeriössä evaluoinnit tavoittavat vain osastot ja yksiköt, jotka ovat suoraan tekemisissä kehityspolitiikan ja kehitysyhteistyön kanssa. Ministeriön johto, johtoryhmät ja ministereiden kabinetit eivät keskustele evaluointien pohjalta.

EVA-11:llä on kaikki eväät jakaa evaluointituloksia UM:n sisällä ja ulkopuolella. EVA-11:n yksikönpäällikkö on laajennetun johtoryhmän jäsen ja voi sitä kautta edistää evaluointien käyttöä tiedolla johtamiseen.

Evaluointiraportit ovat raskaita, ja niiden tyyli karkottaa lukijat. Evaluointien hyödyntämisen esteeksi tulee myös omaksumiskyvyn rajat ministeriössä, jos useita evaluointeja on käynnissä samanaikaisesti.

Nykykäytännön mukaan evaluointiraportteja tarkastellaan lähinnä vain suositusten pohjalta. Raportit sisältävät kuitenkin paljon arvokasta tietoa ja siksi olisi tärkeä pidentää evaluointien elinkaarta nostamalla kiinnostavia aiheita keskustelufoorumeilla, muistuttamalla laajasti ihmisiä evaluointien tuloksista ja mistä evaluoinnit löytyvät.

Myös aiemmat evaluoinnit ja tarkastelut ovat nostaneet esiin tarpeen vahvistaa evaluoinneista oppimista ministeriössä. Oppiminen on toistaiseksi rajoittunut ohjeistojen, työkalujen ja koulutusmateriaalien tuottamiseen. Evaluoinneista oppiminen edellyttää sellaisen lähestymistavan omaksumista, jossa oppimisprosessit ovat fasilitoituja, tiimipohjaisia ja tulevaisuuteen katsovia.

Evaluoinnit ovat hyödyllisiä vain, jos ne valmistuvat oikea-aikaisesti ja vastaavat tiedon tarpeisiin. EVA-11 on teettänyt evaluointien lisäksi joitakin tarkasteluja, mitä on kiitetty. Evaluointisuunnitelman joustavuus on jatkossa yksi tekijä, jolla voi vastata nopeasti nouseviin tarpeisiin.

Johdonvastineprosessi ja siihen liittyvä ohjeistus on parantunut viime vuosien aikana. Prosessissa on kuitenkin ongelmia, jotka haittaavat toiminnan kehittämistä evaluointien pohjalta. Prosessi ei sovi sellaisten suositusten käsittelyyn, joiden vaikutus ulottuu koko ministeriöön. Prosessi edistää tilivelvollisuutta, mutta ei palvele oppimista tai tietojohtamista eikä kerro juurikaan mitä myönteisiä muutoksia on saatu aikaan. Ministeriön rakenteellisia tai institutionaalisia ongelmia ei voi johdonvastineessa käsitellä, koska johdonvastine valmistellaan osastojen tasolla.

4. Suositukset

1) Sitouta uudet ministerit ja heidän kabinettinsa ja lisää kiinnostusta keskitettyihin evaluointeihin ja niiden merkittävyyteen

2) Jaa evaluointitietoa jatkuvalla syötteellä ministereiden kabineteille, johtoryhmille ja kehityspoliittiselle ohjausryhmälle.

3) Käynnistä johdonvastineprosessin uudistaminen

4) Muuta tiedonjakamisen tapoja siten, että se palvelee eri sidosryhmien erilaisia tarpeita

5) Keskity kollektiivisiin oppimismenetelmiin ja osallista UM:n ihmiset

6) Lisää joustavuutta evaluointien suunnitteluun ja monenlaisia evaluointeja

Sammanfattning

1. Översynens syfte, mål och metod

Översyn av användningen och nyttan av centraliserade utvärderingar har gjorts på uppdrag av utrikesministeriets (UM) enhet för utvärdering av utvecklingssamarbetet (EVA-11).

Syftet med översynen är att stärka användningen av utvärderingar för beslutsfattande, lärandet och ansvarighet samt att främja användningen och nyttighet av centraliserade utvärderingar. Därmed bidrar översynen till förbättring och genomförande av utvecklingspolitik och samarbete. De specifika målen är:

- utreda huvudfaktorer som främjar eller förhindrar användningen av utvärderingsresultat och verkställande av rekommendationer både inom UM och med finländska intressegrupper;

- Att identifiera nyckelteman och/eller frågor av strategisk relevans för MFA för vilka det verkar finnas viktiga luckor i genomförandet av rekommendationerna; och

- Att ge kortfattade slutsatser och rekommendationer som hjälper MFA att öka användningen av utvärderingar och genomförandet av rekommendationer.

Översynen baserar sig på ett **team- och samskapande synsätt med betoning på framtiden**. Teamet organiserade under november-december 2022, totalt 36 möten i form av faciliterade workshops, fokusgruppsdiskussioner och individuella eller par diskussion. Informationen erhölls av 62 personer vid utrikesministeriet, Finlands beskickningar, utvecklingspolitiska kommissionen, riksdagen, statens revisionsverk, finska civilsamhällets organisationer, organisationer inom den privata sektorn och utvärderingsexperter. Utvärderingsrapporter, ledningens respons och uppföljningsrapporteringen av ledningens respons samt årsplaner och årsrapporter av utvärdering är några av de studerade dokumenten.

Översynen fokuserar på de centraliserade utvärderingar som beställdes av EVA-11. 23 utvärderingar och tre översyner genomfördes under 2015–2022. De varierar sig från utvärderingar av landstrategier och instrument av utvecklingssamarbete till tematiska och metautvärderingar.

2. Sammanfattande svar på översynens frågeställningar

Översynsfråga 1: Vad är förändringskraften i de centraliserade utvärderingarna?

De centraliserade utvärderingarna har endast haft en liten inverkan på den politiska dialogen. Utöver den utvecklingspolitiska övervalsredogörelsen, som blev färdig 2021, och två utvecklingspolitiska resultatrapporter (2018 och 2022) delade ministeriets personal inga andra exempel på användningen av evalueringar i den politiska dialogen.

Inom UM diskuteras utvärderingar på utvecklingspolitiska styrgruppen och vid avdelningar och enheter som genomför utvecklingspolitik och utvecklingssamarbete. EVA-11 utarbetar sin årsrapport och den går till en utvidgad ledningsgrupp, men gruppen diskuterar inte enskilda evalueringar.

Detsamma gäller utvecklingspolitiska kommissionen och riksdagen och dess utrikesutskottet. Ministrarnas kabinetter får information om genomförda utvärderingar, men den diskuteras inte nödvändigtvis med dem.

Koherens mellan utrikespolitik, säkerhetspolitik, handelspolitik och utvecklingspolitik har endast behandlats i ett par utvärderingar. Utvecklingspolitiken är en del av utrikes- och säkerhetspolitiken, så samstämmigheten i politiken kan förbättras genom att olika politikområden integreras mer systematiskt i utvärderingar.

Vid översynen framgick det mycket positiva bevis att utvärderingarna har påverkat riktlinjer, anvisningar och processer för utvecklingssamarbete. De verkställande enheterna och avdelningarna har betjänats av sådana evalueringar som har ett klart och praktiskt tema. Tidpunkten för utvärderingar togs ofta upp i diskussionerna. De bästa exemplen på de positiva effekterna av utvärdering är de där utvärdering har stött pågående beslutsfattande eller genomförandet av processer. Till exempel finns utvärderingar som fokuserar på landprogram och strategier och resultatstyrning har varit sådana.

Metautvärderingarna har bidragit till utvecklingen av en utvärderingsmanual, verktyg och utbildning. Från finska utvecklingspolitikens synpunkt centraliserade utvärderingar har inte varit transformativa från 2015 till 2022.

Översynsfråga 2: Hur används centraliserade utvärderingar och hur nyttiga är de?

Användningen av utvärderingar i utrikesministeriet och annanstans varierar mycket. Aktiva användare finns på avdelningarna och enheterna, men ju högre hierarkin är, desto värre utvärderingar används.

I utrikesministeriet har utvärderingar bidragit till riktlinjer till exempel för landsprogramming, trippel nexus, samarbetet med civilsamhällesorganisationer och genomgående mål. Utvärderingar har bidragit till metoder bland andra resultatstyrning, riskhantering och rättighetsperspektiv.

Flera finländska civilsamhällesorganisationer utnyttjar utvärderingarna för att utveckla sin egna organisation, lära sig och påverka. Utvecklingspolitiska kommissionens medlemmar och sekretariatet vet om utvärderingarna och uppskattar informationen i rapporterna, men de använder sig inte av dem för närvarande. Inom privata organisationer finns det användare sporadiskt.

Syftet med ledningens responsprocess är att säkerställa att utvärderingsresultaten utnyttjas i UM. Processen har blivit mer strukturerad under de senaste åren men översynen avslöjade ett antal brister. Processen styr i första hand utarbetandet av rekommendationer, men betydligt mindre till att utnyttja övrig utvärdering, nämligen resultaten och slutsatserna.

Ledningens responsmatriser innehåller mycket information och fylls i enligt anvisningarna av EVA-11. Trots det anger de inte alltid explicit om rekommendationerna är överenskomna eller inte. Cirka tio procent av rekommendationerna från utvärderingarna har helt förkastats i processen, i vissa fall utan någon tydlig motivering. Det är nödvändigt att förbättra innehållet i åtgärderna, tidtabellen och ansvarsfrågor.

Enligt uppföljningsrapporterna av ledningens respons har en stor majoritet av överenskomna åtgärder genomförts inom den fastställda tidsramen på två år. Detta tyder på ett stort engagemang för att genomföra åtgärderna. Tvåårsperioden anses dock vara för lång och nuvarande

rapporteringsmatriser innehåller ingen analys av effekten av genomförda åtgärder eller lärande. Rapporterna ger inte information om vad som har förändrats till följd av utvärderingen.

Översynsfråga 3: Vilka är de typiska faktorerna som underlättar eller hindrar upptagandet av centraliserade utvärderingar?

Översynen uppskattar hur mycket EVA-11 har satsat på att förbättra utvärderingsprocessen under översynsperioden. Utvecklingen syns i de verktyg, riktlinjer och utbildning som har gjorts till både UM:s interna and externa deltagare.

EVA-11 identifierar utvärderingar med ett inkluderande begrepp, men i sista hand beslutar EVA-11 vad som evalueras. Tidigare har EVA-11 endast beställt stora utvärderingar med långa tidsperioder (ungefär två år från idé till resultat). De snabba översyner på några månader inleddes i 2021, och ministeriets personal har ansett att de är nyttiga.

EVA-11 har lagt ut genomförandet av utvärderingar Ramtalet om tjänster inom utvärdering av utvecklingssamarbetet (*Evaluation Management Services, EMS*). Det har säkerställt att utvärderingarna görs standardiserade och av jämn kvalitet.

Vid diskussioner uppskattades många öppen kommunikation och möjligheter att delta antingen som informanter eller delta referensgrupper. En workshop där de preliminära resultaten, slutsatserna och rekommendationerna diskuteras ansågs vara nyttig.

När en utvärdering pågår ligger ägandet hos EVA-11 och utvärderingsteamet. Ägandet och ansvaret för genomförandet av utvärderingsrekommendationer överförs till avdelningen eller enheten när rapporten är klar ock beslutsprocessen inleds. Ibland har det varit utmanande att överföra ägande.

Användarna av utvärderingar väntar sig att rapporterna förbättras, att informationen är lättillgänglig och att få skräddarsydd information, till exempel förhandsinformation, några huvudkommentarer, ensidiga sammanfattningar osv. Rapporterna innehåller mycket nyttig information som inte åldras snabbt, så information skulle kunna utnyttjas också på annat sätt än att förbereda ledningens respons.

Den nuvarande ledningens responsprocess i sin nuvarande form fungerar bra när de ärenden som behandlas är på en sådan nivå att respektive avdelningar/enheter och den utvecklingspolitiska styrgruppen har ett klart mandat. Processen fungerar dock inte när rekommendationer gäller UM:s människor frågor. I hälften av utvärderingarna under översynsperiod gällde rekommendationer ministeriets underbemanning eller människornas kompetensutveckling och i många andra utvärderingsrapporter fanns resultat eller slutsatser om dem. Ministeriets administrativa tjänster agerar inte utifrån utvärderingsrapporten och inte heller UM:s ledning har agerat i frågan. Att personalfrågan inte är avgjord utgör en betydande risk för Finlands utvecklingspolitik och samarbete.

Översynsfråga 4: Hur skulle MFA bättre utnyttjainformationen och lärdomarna från utvärderingar som ett verktyg för kunskapsbaserad ledning?

Under översynsperioden 2015–2022 utnyttjades utvärderingarna kunskapsbaserat beslutsfattande på avdelnings- och enhetsnivå vid UM, men utvärderingarnas potential att stödja kunskapsbaserad ledning har inte uppnåtts genom hela UM, inklusive ministeriets ledning. Resultaten används inte heller av de politiska beslutsfattarna.

Användningen av utvärderingar för lärande är snarare beroende av individerna än av organisations lärande. I översynen framkomgott om bevis hur individer använder och lär sig av utvärderingarna. Inlärningen utmanas av informations inflykten eftersom utvärderingarna inte är de enda viktiga informationsprodukterna. Ministeriets kontext, så som personalrotation, byte av specialtjänstepersoner och otillräcklig personalyrka i förhållande till uppgifterna) har också en negativ påverkan på lärandet från utvärderingar och spridningen av information.

Att delta i utvärderingsreferensgrupper har visat sig vara en fantastik väg till teaminglärning. Inspirerande idéer och samarbete har lett till att deltagarna genomför förändringar redan under utvärderingen. Det sades inte samma för arbetsgrupperna som förbereder ledningsresponsen; det kom inte fram att dessa grupper skulle ha främjat insikter och lärande. Tills vidare betraktas utvärderingar som inlärning endast vid vissa avdelningar och enheter inom UM, men resten av organisationen känner inte till denna tanke.

Översynen fann flera problem med utnyttjandet av utvärderingar. Många som intervjuades berättade om arbetsbelastningen och hektiska arbetsdagarna, så det finns inte mycket arbetstid för inlärning, reflektering och behandling av ny information, vilket hindrar att utvärderingsresultaten utnyttjas. Den kunskap som utvärderingarna generar överskrider således inte enhetsgränserna vid UM. Det kom också fram att utvärderingarna är svårt att hitta i UM:s informationssystem.

Utvärderingarna har potential för lärande och informationsförädling. EVA-11-teamet har en vision om centraliserade utvärderingar som en betydande del av främjandet av kunskapsbaserad ledning i ministeriet. Teamet har djup kännedom om tidigare och kommande utvärderingar och har möjlighet att påverka beslutsfattarna, eftersom EVA-11 är presenterad i alla viktiga grupper. EVA-11 har också nyttig erfarenhet av sådana kollektiva inlärningslösningar som kan användas till att vårda inlärning och förädling av information tillsammans.

3. Slutsatser

Under översynsperioden har EVA-11 strävat efter för att förbättra utvärderingar och utvärderingsprocessen genom att involvera människor mer, utveckla anvisningar och verktyg samt informationsprodukter. Detta utvecklingsarbete har behövts och översynen uppskattar och anser att dessa åtgärder är ändamålsenliga för att främja lärande och kunskapsbaserad ledning i ministeriet.

Översynen har identifierat flera områden som behöver förbättras. Det är viktigt att komma ihåg att EVA-11 endast delvis har användning och användbarheten i sina händer. Hur nyttiga utvärderingarna är beror på vilka avdelningar och enheter som ska äga resultaten och utnyttja dem i praktiken.

Centraliserade utvärderingarna har påverkat och förändrat riktlinjer, anvisningar och utvecklingssamarbetets processer. Effekterna på den politiska dialogen är begränsade, även på utvecklingspolitiken. Utvärderingarna når inte de politiska beslutsfattarna i parlamentet och i utvecklingspolitiska kommissionen. Medlemmarna i kommissionen inser värdet och kvaliteten på utvärderingarna, men de används inte. Vid UM når utvärderingarna endast de avdelningarna och enheterna som är involverade i utvecklingspolitik och samarbete. Ministeriets ledning, ledningsgrupper och ministeriets kabinetter diskuterar inte utifrån utvärderingar.

EVA-11 har all förmåga för att dela ut utvärderingsresultat både inom och utanför MFA. Enheten är till exempel medlem i den utvidgade ledningsgrupp och kan främja användningen av utvärderingar för kunskapsbaserat beslutsfattande.

Utvärderingsrapporterna är tunga och typen av rapporter utvisar läsarna. Ett hinder för utnyttjandet är också gränserna för förmågan att tillägna sig vid ministeriet, om flera utvärderingar pågår samtidigt.

Enligt nuvarande praxis granskas utvärderingsrapporter endast utifrån rekommendationer. Utvärderingarna innehåller dock mycket värdefull information och därför är det viktigt att förlänga utvärderingens livscykel genom att lyfta fram intressanta teman på relevanta diskussionsforum genom att i stor utsträckning påminna folk om resultaten av utvärderingarna och var utvärderingarna finns.

Också förra utvärderingar och övertyg har väckt uppmärksamhet till behov att förstärka inlärning från utvärderingar i UM. Än så länge aktiviteter har koncentrerat på utveckling av riktlinjer, verktyg och träning materials. Inlärning från utvärderingar kräver antagandet av inlärningsmetoder som innehåller underlättade inlärningsprocesser, är team baserad och med feedforward fokus.

Utvärderingar är nyttiga bara om de blir färdiga vid rätt tidpunkt och svarat mot informationsbehovet. EVA-11 har utöver utvärdering också låtit göra vissa översyner, vilka har berömts. I fortsättningen är flexibiliteten i evakueringsplanen en faktor som gör det möjligt att snabbt svara mot växande behov.

Ledningens responsprocess och tillhöriga anvisningar har förbättrats under de senaste åren. Det finns dock problem i processen som försvårar utvecklingen av verksamheten utifrån evalueringar i UM. Processen är inte lämplig för behandling av sådana rekommendationer vars verkan överskrider de institutionella gränserna vidministeriet. Processen främjar ansvarsskyldighet, men tjänar inte inlärning eller kunskapsbaserad ledning och berättar inte i någon större utsträckning vilka positiva förändringar som har åstadkommits. Strukturella eller institutionella frågor kan inte tas upp eftersom ledningens responsprocess bara bereds på avdelningsnivå.

4. Rekommendationer

1) Engagera de nya statsrådsministrar och kabinetter och öka deras intresse för betydelsen av de centraliserade utvärderingarna.

2) Ständigt dela utvärderingsinformation med ministerkabinetter och ledningsteam och i utvecklingspolitiska kommissionen.

3) Inled en process för att tänka om och revidera ledningens responsprocess

4) Förbättra informationsutbytet om utvärderingarna för att svara mot de olika intressenters olika behov

5) Prioritera kollektiva lärandeorienterade metoder i syfte att engagera MFA:s intressenter

6) Öka flexibiliteten och mångfalden av utvärderingar

Summary

1. Purpose, objectives and methodology of the review

The Review of the Use and Utility of Centralised Evaluations was commissioned by the Development Evaluation Unit (EVA-11) of the Ministry for Foreign Affairs (MFA).

The purpose of the review is to strengthen the use of evaluation evidence for decision making, learning and accountability and to the enhance the utility of evaluations. The review contributes to improvement and implementation of development policy and cooperation. Its specific objectives are:

- To identify main factors that enable or hinder the uptake of evaluation results and implementation of recommendations both in the MFA as well as among other stakeholders in Finland;

- To identify key themes and/or issues of strategic relevance to the MFA for which there seems to be important gaps in implementation of recommendations; and

- To provide concise conclusions and recommendations that help the MFA to step up the use of evaluations and implementation of recommendations.

The review is based on a team and co-creation approach with future orientation. The team organised during November-December 2022, a total of 36 data collection sessions consisting of facilitated workshops, focus group discussions and individual or pair discussions. The persons consulted (62 in total) represent the MFA, the Finnish Embassies, the Development Policy Committee, the Parliament, National Audit Office, Finnish CSOs, private sector organisations, and evaluation experts. Evaluation reports, management responses and their follow-up reports, annual evaluation plans, and annual reports of development evaluation are among documents studied.

The review focuses on the centralised evaluations that are commissioned by the EVA-11. 23 evaluations and three reviews or assessments were completed during 2015-2022. They range from country strategy and aid instrument evaluations to sector policy and thematic evaluations, and meta-evaluations.

2. Summary answers to the review questions

Review Question 1: How transformative are centralised evaluations?

The centralised evaluations have only modest influence on policy dialogue. Apart from the preparation of the 2021 Report on Development Policy Across Parliamentary Terms and the 2018 and 2022 development policy result reports, the people working in the MFA were not able to provide much information about using the evaluations in policy dialogue.

In the Ministry, discussion on evaluations remains at the level of departments and units that engage in development policy and development cooperation. The EVA-11 Annual Report is brought

to the Extended Management Team of the Ministry, but no individual centralised evaluations are discussed there. The same applies to the Development Policy Committee and the Parliament and Foreign Affairs Committee. The Offices of the Ministers receive evaluation information, but it is not discussed with them.

Coherence between foreign, security, trade and development policies has only been addressed in a couple of evaluations. Since development policy is an integral part of Finnish foreign and security policy, also opportunities to improve policy coherence at the MFA through addressing foreign, security and trade policy in the future evaluations would be important.

The review found plenty of positive evidence about how the evaluations have influenced development cooperation policies, guidelines and processes in the MFA. Practical themes have been relevant to operational units and department levels. Timing of evaluations was raised often in discussions. The best examples that have had positive influence are those that have fed into ongoing decision making and implementation. Amongst such are the evaluations focusing on country strategies and programmes, and Results-Based Management.

The meta-evaluations have fed into the development of the Evaluation Manual, tools and training contents. From Finnish development policy viewpoint centralised evaluations have not been transformative during 2015-2022.

Review Question 2: How are centralised evaluations used and how useful they are?

The degree which the stakeholders both inside and outside MFA use the centralised evaluations varies. Active users of evaluations are found at the department and unit level. The higher the post in the decision making chain in the Ministry, the less evaluations are used.

In the MFA the evaluations have fed, for example, into the guidelines for country programming, triple nexus, CSO cooperation, and Cross-Cutting Objectives. Among others, the evaluations have contributed to development of approaches like Results-Based Management, risk management and Human Rights Based Approach.

Several Finnish Civil Society Organisations utilise the evaluations in the development of their own organisation, learning and advocacy work. The Development Policy Committee members and secretariat are aware of the evaluations and appreciate the high information value of the reports but are not using them at present. There are random users of evaluations in the private sector organisations.

The management response process is supposed to ensure that evaluation results are utilised. It has become more structured during the period review, but a number of shortcomings were identified. The process directs the focus on recommendations and much less on the other valuable information provided in the evaluation reports, namely the findings and conclusions.

Although the management response follow-up tables are informative and are keeping with the structure and guidelines prepared by the EVA-11, they do not always explicitly state if the recommendations are agreed or not. About ten percent of the recommendations made in the evaluations were fully rejected in the process, in some cases without clear justifications. There is room for improving the content of actions, definite deadlines, and responsibilities.

According to the follow-up reports of management responses, a vast majority of agreed actions have been implemented within the stipulated time frame of two years. This indicates a high commitment to implementing the actions. The two-year period, however, is considered too long and the present reporting tables do not contain any analysis about the effect of implemented actions or learning taking place. The reports do not give information regarding what are the positive changes as a result of the evaluation.

Review Question 3: What are the typical factors facilitating or hampering the uptake of centralised evaluations?

The review appreciates that the EVA-11 has put a lot of emphasis on the improvement of evaluation process during the period of review. This is reflected in the tools, guidelines and training provided to MFA internal and external participants.

The EVA-11 applies a participatory process of identifying ideas for evaluations, but the final selection of the evaluation topics is in the hands of the EVA-11. In the past the EVA-11 has commissioned only large evaluations with long time span (approximately two years from the idea to availability of results). Reviews that can be completed in a matter of months were introduced to the portfolio in 2021 and are considered very useful by the people in the Ministry.

The EVA-11 has an outsourcing arrangement, Evaluation Management Services (EMS) contract for conducting the evaluations. The EMS process has evolved into conducting evaluations in a standardised manner and with even quality.

During evaluations, open communication and opportunities to participate either as informants or reference group members was appreciated by many. The Findings, Conclusions and Recommendations workshops that are organised in every evaluation were considered useful.

When an evaluation is ongoing, the EVA-11 and the evaluation team jointly own the process. Changing the ownership from the EVA-11 to the department/unit leading the management response process has been challenging at times.

Evaluation users expect improvements in style of reports, easy access to evaluations, and receiving tailor-made information, including advance information, short lists, one pager briefs, etc. The reports contain plenty of useful information that remains valid for many years, and the information could be utilised not just in the preparation of the management response.

The present management response process works well in addressing operational and strategic recommendations that clearly fall within the mandate respective departments/units. However, recommendations dealing with people issues in the MFA and the Embassies are not among these. Half of the evaluations made recommendations about the understaffing or competence development needs, and many others had findings or conclusions on them. The Administrative Services of the Ministry do not act independently on any evaluation report and the MFA leadership has not acted on these issues. Not addressing them is a significant risk for Finnish development policy and cooperation.

Review Question 4: How could the MFA better capture and retain knowledge and learning from evaluations as one tool for knowledge-based management?

The evaluations benefit knowledge-based decision making at the department and unit level but the MFA leadership, namely Management Teams and Offices of the Ministers do not use the evaluation results. The results are not used by the political decision makers either. The potential of evaluations for supporting knowledge-based decisions at the level of Ministry leaders is not reached.

Use of evaluation results rests more on individuals than on the organisation of the Ministry. There is abundance of evidence from individual efforts to generate knowledge and learning from the evaluations. The learning is challenged by information overflow as the evaluation reports are not the only important knowledge products. Internal contextual factors of the Ministry (rotation of post holders, continuous change of special public servants, understaffing) also influence retaining knowledge and learning from evaluations negatively.

Participating in evaluation reference groups can be a team learning journey at best. Inspiring ideas and collaboration make participants to implement changes already during the evaluation. The working groups preparing the management responses were not mentioned in this context; there was no talk about those groups provoking insights and learning. So far only a few units and departments embrace the idea for using evaluations for learning, the rest of the organisation has not quite grasped it.

Several problems in retaining knowledge from evaluations were identified. For example, it is evident that time for learning, reflection and retaining knowledge is in short supply as many people spoke of their workload and hectic workdays and how it hinders the use of evaluation results for any purpose. The evaluations are unable to transfer knowledge they generate across boundaries in the organisations. The evaluations are also not easily found in the MFA information systems.

Potential and ideas for capturing knowledge and learning from the evaluations exist. The EVA-11 team has a vision of centralised evaluations as an essential part of contributing to knowledge-based management in the Ministry. Equipped with both the profound knowledge of the past and planned evaluations the EVA-11 has the potential to influence the decision makers because it has a seat in all relevant groups. The EVA-11 has also generated useful experience from collective learning solutions that can be used for nurturing learning and generating knowledge from evaluations.

3. Conclusions

During the period of review, the EVA-11 has worked towards improving evaluations and the evaluation process by increasing the engagement of stakeholders, by developing formats and tools and by developing information products and materials. The review supports these measures fully and considers them relevant for using the evaluations for learning and knowledge-based management in the Ministry.

The review has identified several areas requiring improvement. Importantly, the use and utility of the evaluations is only partly in the hands of the EVA-11. Ultimately the utility of the evaluations depends on departments and units that are expected to embrace the results and apply them in practice.

The centralised evaluations have been influential in terms of changing policies, guidelines and processes of development cooperation. The influence on policy dialogue, is limited, even in development policy. The evaluations do not reach the political decision makers in the Parliament and in the Development Policy Committee (DPC). The members of the DPC recognise the value and quality of the evaluations, but do not use them. Importantly, in the Ministry the evaluations only reach the level of Director Generals of Departments and Directors of Units involved with development policy and cooperation. The Ministry leadership, the Management Teams and the Offices of the Ministers, do not discuss the evaluations.

The EVA-11 has all the power to improve sharing of evaluation results inside and outside the MFA. For example, the Unit is a member of the Extended Management Team and can advocate for using evaluation knowledge in decision making.

The use of evaluations is hampered by the volume and style of evaluation reports. Use of evaluations is also hampered by absorption capacity in the Ministry when many evaluations are ongoing at the same time.

At present the evaluation results are approached from the angle of recommendations only. The evaluations contain a considerable amount of valuable information and knowledge, and it would be important to extend the shelf life of a report by raising interesting topics for relevant discussion fora and for reminding people of the evaluations and where to find them.

Also previous evaluations and reviews have raised the need to strengthen learning from evaluations. So far, the activities have focused on development of guidelines, tools and training materials. Learning from evaluations calls for adopting a learning approach where learning processes are facilitated, team based and with a feed-forward focus.

Evaluations can be useful only if they are available in a timely manner and address direct information needs. The steps the EVA-11 has adopted in diversifying the types of evaluations and commissioning more reviews is welcome. Having more flexibility in the Annual Evaluation Plan would be helpful in responding to rapidly arising needs.

The management response process and its guidelines have been improved. But the process has several problems that hamper its usefulness. The process is not fit to address recommendations that transgress the institutional boundaries in the Ministry. The process contributes to accountability but does not instigate learning or knowledge-based management and evidence of change is almost non-existent. No structural or institutional issues can be addressed because the management response process only reaches the department level.

4. Recommendations

The review has six recommendations.

1) Engage the new Ministers and people in their Offices and raise their interest in the significance of the centralised evaluations.

2) Constantly share evaluation information and knowledge with the Offices of the Ministers, Management Teams, and in the Development Policy Steering Group.

3) Initiate a process of rethinking and revising the management response process.

4) Improve on the information sharing about the evaluations to respond the varied needs of different stakeholders

5) Prioritise collective learning oriented methods in engaging MFA stakeholders

6) Increase flexibility and diversity of evaluations

1 Introduction

This report presents the findings, conclusions and recommendations of an independent review, commissioned by the Development Evaluation Unit (EVA-11) of the Finnish Ministry for Foreign Affairs (MFA). It assesses the different types of evaluations, meta-evaluations and reviews that the EVA-11 has carried out in 2015-2022. The review contributes to the analysis of their use and utility as drivers for change and tools for learning and knowledge-based management.

1.1 This review

The review covers the period from 2015 to 2022. The Terms of Reference (ToR) lists as the main intended users of the review the EVA-11, the Development Policy Steering Group, the departments managing development policy and cooperation, and senior management of the MFA. Other Finnish stakeholders outside the MFA are expected to be interested in the results of the review as well, namely the Development Policy Committee, Parliament, Civil Society Organisations (CSOs) and private sector entities (Terms of Reference, Annex 2).

The **purpose of the review** is to strengthen the use of evaluation evidence for decision making, learning and accountability and to enhance the use and utility of centralised evaluations. The review contributes to improvement and implementation of development policy and cooperation. Its specific objectives are:

- To identify main factors that enable or hinder the uptake of evaluation results and implementation of recommendations both in the MFA as well as among other stakeholders in Finland;

- To identify key themes and/or issues of strategic relevance to the MFA for which there seems to be important gaps in implementation of recommendations; and

- To provide concise conclusions and recommendations that help the MFA to step up the use of evaluations and implementation of recommendations.

The review is expected to provide answers to four **main review questions (RQ)**:

- RQ 1. How transformative are centralised evaluations?

- RQ 2. How are centralised evaluations used and how useful they are?

- RQ 3. What are the typical factors facilitating or hampering the uptake of centralised evaluations?

- RQ 4. How could the MFA better capture and retain knowledge and learning from evaluations as one tool for knowledge-based management?

Terminologies used in the review

In this review the **term "MFA leadership"** is defined to consist of the top of the Ministry meaning both people in political and civil servant roles. To be more precise, the MFA leadership consists of Ministers and the people in their Offices, State Secretary and Under-Secretaries, and the MFA Management Teams.

This review does not use the collective term of "human resources" because the resource talk overlooks the respect for human being. Therefore, the **review talks of people** reflecting the change in thinking.

1.2 Methodology

The approach and methodologies of the review were agreed with the EVA-11 during the inception in October 2022. This is a review with a short duration (four months) and a small team (two experts). It is based on a team and co-creation approach and has future orientation. Documents provide valuable data about the evaluations, but they only speak of the past. The discussions with the MFA people and other stakeholders provide information about the present use and utility of evaluations and the potential for future which underlines the significance of discussions and dialogue in the methodology.

During the inception, representatives of main stakeholder groups that were expected to be among the potential users of the evaluations were identified. Subsequently, the team organised 36 data collection sessions consisting of facilitated workshops, focus group discussions and discussions with individuals or in pairs. Few persons shared their views over email. In total, 62 persons engaged with the review. They represent the Ministry for Foreign Affairs, the Finnish Embassies, the Development Policy Committee, the Parliament, National Audit Office, Finnish Civil Society Organisations, private sector organisations, and evaluation experts. All discussions were noted down. The findings of the review are grounded on the views, perceptions and experiences shared by all these people.

Review methods have had an emphasis on future and co-creation.

Documents are an important source of supplementary data. The documents studied consist of reports depicting different steps in the evaluation process from mandate of development evaluation to planning evaluations, managing and conducting evaluations, and finally addressing recommendations of the evaluations. Among the key documents studied are the Ministry's Development Evaluation Norm and Rules of Procedure. Annual Evaluation Plans, Annual Reports of Development Evaluation, and several guidelines that the EVA-11 has prepared for decentralised evaluations were studied. Reports of individual evaluations together with respective management responses and their follow-up reporting have been looked at. Also other MFA documents related to development policy and development cooperation in the MFA were reviewed.

The review team conducted joint data analysis during the first half of January. In the analysis, each stream of evidence (interview notes and documents) was systematically analysed against the Review Matrix questions. Data triangulation was a two-step process, i.e. qualitative discussion

data was separately collated and analysed against the questions in the Review Matrix. Key words were used to classify responses and used to identify narrative patterns. Then, a systematic content analysis was conducted for the documents, and findings compared with the discussion findings. As Chapter 3 shows, sometimes the views presented by the stakeholders agreed with each other and sometimes they contradict each other. Same applies to findings substantiated by the documents.

The analysis was boosted by discussions with the MFA representatives in the joint Findings, Conclusions and Recommendations (FCR) workshop in January 2023. The EVA-11 team together with the representatives from the Department for Africa and Middle East, Department for the Americas and Asia and Department for Development Policy participated in the workshop.

The Review team is presented in Annex 1. The approach, methodology and limitations of the review are discussed in Annex 3, references are in Annex 4 and list of evaluation reports, management responses and follow-up reports studied is in Annex 5.

1.3   Structure of the report

Chapter 2 presents the background of the review, Chapter 3 the findings on the review questions and Chapter 4 the conclusions of the review team. It is followed by recommendations in Chapter 5.

2   Background and Context

2.1   Background

The mandate of the EVA-11

The overall framework of evaluations in the Ministry for Foreign Affairs is set in **the evaluation norm** (MFA 2015a). The norm provides legal basis of evaluation of development policy and development cooperation and guides all evaluation function of these two in Finland's foreign affairs. The norm also defines the purposes of development evaluation: in the Ministry development evaluation is conducted for accountability, organisation-wide learning, improvement of quality, transparency and openness (Box 1).

> **Centralised evaluations are managed by the EVA-11.**

Box 1. Development evaluation for many purposes

"Evaluation of development policy and cooperation (development evaluation) is a part of the responsibilities (accountability) for relevance, efficiency, effectiveness, and impact and their evaluation that is set for the MFA by the State Budget Act and State Budget Decree. Development evaluation serves the organization-wide learning from development cooperation experiences, as well as the improvement of quality by producing independent and impartial evidence-based knowledge of the activities. It is also an essential part of transparency and openness of development policies."

Source: Development Evaluation Norm (MFA 2015a)

The evaluations commissioned by the MFA are divided into **two types of evaluations** based on management responsibility and scope (MFA 2015a). Centralised evaluations are managed by the EVA-11 and are characterised as being comprehensive and strategic. Examples of such evaluations are evaluations of policies, country strategies, financing instruments, processes, results- , theme- or sector-based programmes or programmes of partner institutions. Decentralised evaluations are managed by other Ministry units and Embassies implementing development cooperation, and they focus on projects and programmes.

According to **the Rules of Procedure for the Ministry** (Decree of the Ministry for Foreign Affairs 550/2008), the EVA-11 is a member of the Development Policy Steering Group and the Extended Management Team of the Ministry. Among the five functions of the Development Policy Evaluation Unit two are of particular relevance for this review, namely function number 1 on ensuring centralised evaluations of development policy and cooperation and function number 4 on disseminating development evaluation results.

Evaluation process

Following the approval of the Evaluation Norm in 2015, **a process for centralised evaluations** has evolved. The process begins from planning the topics for all centralised evaluations and moves then into planning and management of an actual evaluation, preparing a management response, and ends with the follow-up report of the management response for each respective evaluation (MFA 2015a). Annex 6 provides a description of the evaluation process.

Evaluations and reviews during 2015-2022

Starting from 2015, a total of 23 evaluations have been completed. The evaluations consist of country strategy evaluations, policy, thematic and sector-based evaluations, evaluations of processes and programmes, evaluations of aid instruments and meta-evaluations of decentralised evaluations. In addition, three other assignments (assessments and reviews) have been completed (Table 1.)

Table 1. Centralised evaluations completed during 2015-2022

TYPE	NAME OF EVALUATION AND YEAR COMPLETED
Country strategy evaluations	Evaluation on Finland's Development Cooperation with Kenya 2007-2013 (2015) Evaluation on Finland's Development Cooperation Country Strategies and Country Strategy Modality (2016) Adapting for Change: Country Strategy Approach in Fragile Contexts (2020) Evaluation on Development Cooperation Carried out by the Department of Russia, Eastern Europe and Central Asia, including the Wider Europe Initiative (WEI), (2021) Evaluation on the Transition Process of Finnish Vietnamese Cooperation in 2008-2020 (2021)
Policy, thematic and sector evaluations	Evaluation on Finland's Development Policy Programmes from Results-Based Management Point of View (2015) Evaluation of Inclusive Education in Finland's Development Cooperation 2004-2013 (2015) Evaluation of Aid for Trade (2016) Evaluation on Improvement of Women's and Girls rights (2018) Evaluation on Forced Displacement and Finnish Development Policy (2019) "How do we Learn, Manage and Make Decisions in Finland's Development Policy and Cooperation". Management of results information and knowledge at the Ministry for Foreign Affairs of Finland (2019) Evaluation of Economic Development, Job Creation and Livelihoods (2021) Evaluation of the Finnish Development Policy influencing in the European Union (2022) Evaluation of Finland's Humanitarian Assistance 2016-2022 (2022) Evaluation of Finnish Development Policy Influencing Activities in Multilateral Organisations (2020) Water as an Entry Point for Peace Mediation. Evaluation on Finnish Water Diplomacy (2021)
Evaluation of processes or programmes	Evaluation of Humanitarian Mine Action (2015) Developmental Evaluation of Business with Impact 2015-2019 (BEAM) Programme (2019)
Evaluations of financing instruments	Programme-based Support Through Finnish Civil Society Organizations (Meta analysis 2017)
Meta-evaluations	Meta-evaluation of project and programme evaluations in 2017-2020 (2022) Meta-evaluation of project and programme evaluations in 2015-2017 (2018) Meta-evaluation of project programme evaluations 2014-2015 (2017) Meta-evaluation of project and programme evaluations in 2012-2014 (2015)
Reviews and assessments	A review of Finland's cooperation supporting peace processes (in Finnish) (2021) Business to Government – a developmental review of the platform pilot (in Finnish) (2021) From Reactivity to Resilience. Assessment of the Response of Finnish Development Policy and Cooperation to the COVID-19 Pandemic (2022)

Source: Review team

Baseline for the review

Use of evaluations for learning and knowledge management has been high on the agenda of the EVA-11 for years. Three recent evaluations and reviews have addressed these issues. They provide both justifications for the review and a baseline for it (see ToR in Annex 2).

According to the **evaluation on management of results information and knowledge** (Palenberg et al. 2019) centralised evaluations, their synthesis in annual reports and performance audits have successfully influenced decision making at the MFA. The evaluations produce useful results, but the Ministry does not fully utilise the comprehensive body of knowledge they generate. A moderate level of influence on both learning and decision making at the MFA was evident. The evaluation considered that centralised evaluations are targeted primarily at senior MFA management and the Finnish government and parliament. The relevance and usefulness to MFA staff in general has remained limited. Among the measures to improve the relevance and usefulness the evaluation recommended that relevant findings and conclusions could be "filtered" and "packaged" into suitable formats for different, more specifically defined audiences in (and beyond) the MFA. Increased staff participation would increase understanding and ownership of evaluation findings and the feasibility of recommendations and strengthen learning during the entire evaluation process.

> **Previous evaluations made recommendations about learning and sharing of evaluation results.**

MFA participated in **the joint peer review of the evaluation units** together with Switzerland and Ireland in 2020 (Peer Review of Four Evaluation Functions 2020). Compared with the other organizations that were part of the Peer Review, centralised evaluations commissioned by the MFA tend to be more ambitious in scope, more complex in methods, and longer in duration. The Peer Review considered that all of these factors drive costs upwards, and also have implications for the utility of the resulting products. In the Peer Review interviews operational managers often characterized the reports commissioned by the EVA-11 as "theoretical", almost certainly reflecting a heavy emphasis on the theory of change, broad context, and evaluation methods front-loaded in the reports. The centralised evaluations are geared toward the broader issues facing the organization which the Peer Review considered reasonable. Recommendations were made towards enhancing the planning phase and management response process and improving communication and dissemination of evaluation results.

The **meta-evaluation of MFA's project and programme evaluations in 2017–2020** (Väth et al. 2022) looked at the use and utility of decentralised evaluations, i.e. evaluations commissioned by other departments and units of MFA and Finnish Embassies. Decentralised evaluations are often used for learning in teams, decision making, planning of follow-up, and management of ongoing interventions. The learning, however, stays more on individual level. A large majority of the evaluations were assessed as satisfactory or very useful. Timing of the evaluations, relevance of recommendations, the extent of recommendations being realistic, the learnings from the evaluation, and the implementation of recommendations were positively assessed.

OECD Review of Evaluation Systems

The 2022 Review of Evaluation Systems (OECD 2023) draws on the experiences of the members and observers of the DAC Network on Development Evaluation (EvalNet). The review identifies

major trends in evaluation systems, covering the role of evaluation, policy and institutional arrangements, evaluation processes and the use of evaluation findings. The OECD Review findings point to areas that are also relevant to this review, for example:

1. Work towards a more holistic understanding of all the evidence that is available within an organisation as part of a broader shift towards learning.

2. Track when and how evaluation findings are used. A common theme within the OECD Review relates to challenges in meeting learning objectives and the subsequent focus on finding new and more effective approaches to increase use.

3. Revisit questions of ownership, as related to partner country engagement.

The OECD Review also talks about the evaluation process itself, about how decisions on which evaluations to undertake are made impact dissemination of findings and use of evaluations. In **evaluation plans** the Review argues that the process which the evaluations are undertaken impacts both the quality and usefulness of the findings. It points to three issues, namely linking evaluations to organisational priorities, aligning selection of evaluation topics to organisational priorities, and consultative approach in designing evaluation plans to ensure utility. The OECD Review also brings forward valuable lessons for evaluation design and conducting evaluations.

> **Review argues that evaluation process impacts both quality and usefulness of the findings.**

The Review draws attention to **using evaluation findings**. In the past, promoting the use of findings has been left until after evaluations are complete, rather than considering learning objectives in the early stages of the evaluation process. It argues that timeliness of evaluation findings is essential to their uptake but there is tension between providing quick evidence that can be used to inform ongoing efforts and ensuring high-quality and credible findings. Absorption capacity of individuals and organisations affects the ability of target audiences to act on findings.

Recent **focus on learning objectives** reflects challenges in systematically using evaluation findings to inform decision making and strengthen overall effectiveness. It indicates that more work can be done to facilitate the use of evaluation findings. The organisations participating in the EvalNet complement full length reports with other messaging, in an effort to better connect with internal and external audiences and facilitate use of results.

> **OECD Review supports increased focus on interactive learning.**

Dissemination to internal audiences often extends beyond simple sharing of knowledge products, with increased focus on interactive learning. While evaluation units continue to share evaluation reports and other knowledge products across their organisations increased emphasis is placed on interactive methods of engagement. For example in France the directorate responsible for international development co-operation organises a series of presentations and workshops with senior-level officials and relevant departments to discuss findings and provide support in responding to evaluation recommendations. This consultative approach helps to ensure buy-in to recommendations, thereby increasing their implementation.

The OECD Review proposes the following **items to mitigate challenges in use and utility of evaluations**:

- Align to organisation strategic priorities

- Engage with senior management and subject-matter experts

- Design ToRs that incorporate learning objectives

- Prepare a communication strategy during the evaluation planning phase

- Share findings with key stakeholders on an ongoing basis

- Formulate concrete and actionable recommendations

- Draft reports in non-technical language

- Use a variety of presentation and dissemination methods

2.2   Context

In this chapter a few important context issues are discussed. The issues have implications on the use, utility and influence of the centralised evaluations.

Rapid change is a challenge for all decision-makers. We live in **a time of disruptive change with many overlapping crises** (e.g. COVID-19, Russian aggression against Ukraine, climate change). The leadership of the Ministry benefits from up-to-date information and evaluations can be one important source of such information. This came repeatedly up in the discussions both with the people working in the Ministry and with the external stakeholders.

Finnish foreign and security policy aims to enhance security and wellbeing through international cooperation. The Government Programme (Programme of Prime Minister Sanna Marin's Government, 2019), the Government Report on Finnish Foreign and Security Policy (Government of Finland 2020) and the Report on Development Policy Extending Across Parliamentary Terms (MFA 2021a) explicitly state that **development policy is an integral part of Finland's foreign and security policy**. This linkage was recognised by the MFA people and the members of the Development Policy Committee.

Development policy is an integral part of Finland's foreign and security policy.

The aid cuts made during the Sipilä Government have long-lasting impacts. The cuts have led into significant changes in the structure of Finland's Official Development Assistance (ODA) and how it is presently managed (see data in Annex 8). Compared to 2014, the share of ODA managed by the Ministry has shrunk from 80 % in 2014 to 59 % in 2021 (MFA 2022a). Disbursements to multilateral development cooperation, European Development Fund, Finland's share of EU's cooperation budget and core support to multilateral organisations have all grown. Bilateral aid took really the brunt of the budget cuts made in 2016. Both the actual disbursements and the share of bilateral aid of total Finnish aid keeps shrinking. This is important from evaluation viewpoint because the bilateral partner countries

and their institutions have been a valuable evidence base to so many evaluations in the past. More and more of bilateral partner countries are fragile and conflict countries. The portfolio of aid is also becoming more diverse with new instruments (e.g. development policy loans and investments) and arrangements (e.g. outsourcing to FINCEED, Finnish Centre of Expertise in Education and Development in the National Board of Education and Rule of Law Centre in the University of Helsinki).

Even though the total disbursements have grown, **disbursements for administrative costs have remained stagnant starting from 2016 onwards**. Also development cooperation personnel were reduced during the Sipilä Government. This translates as less people managing a growing volume of aid (National Audit Office 2021). Karhu & Lanki (2022) argue that this is potentially endangering to both quantity and quality of Finnish development.

3 Findings

Discussion on findings begins with use and usefulness of centralised evaluations (RQ2) and is followed with factors facilitating or hampering the uptake of centralised evaluations (RQ3) and how transformative they are (RQ1). The chapter ends with discussion on how the MFA could retain learning from evaluations and foster knowledge-based management (RQ4).

3.1 Use and usefulness of centralised evaluations

This chapter looks at the **Review Question 2: How are centralised evaluations used and how useful they are?** The sub-question on the evaluation processes relates to factors that facilitate or hamper the use and taking up of evaluations, and it is discussed in chapter 3.2.

Key findings

The evaluations are of high quality, they contain useful information and are considered valuable and well-justified sources of information.

People in the Ministry use the centralised evaluations but to a varying degree. Active individuals who use evaluations are found at the department and unit level. The evaluations do not reach the Management Teams or the Offices of the Ministers.

Finnish CSOs apply the information from the evaluations in their own functions. People in the Development Policy Committee appreciate the high information value of the reports, but do not use them.

Improvement is observed in the clarity and focus of the recommendations made in the evaluation reports.

Recommendations related to the number of people or their competence in the Ministry and/or at the Embassies came up in more than half of the evaluations completed during the period of review.

Management response process is expected to ensure that evaluation results are utilised. Although it has become more structured, the process still contains a number of shortcomings. It directs the focus of evaluation users on evaluation recommendations with less attention on findings and conclusions.

Management responses do not always explicitly state if the recommendations are agreed with or not.

A vast majority of actions agreed in the management responses have been implemented. The two-year period is considered too long. The follow-up tables do not contain any analysis about the effect of implemented actions or learning taking place.

3.1.1 Perceptions of usefulness and actual use

This chapter talks first about the views of the MFA and the Embassy staff regarding their motivations and ways how they use evaluations and how useful they find them. It then moves to the views of the other Finnish stakeholders. The extent that evaluation process supports use of evaluation results is merged with factors that facilitate or hamper use of evaluation results and is discussed in chapter 3.2.1.

Views from the MFA and the Finnish Embassies

The review engaged with more than 30 people with current or recent working experience at the MFA or in the Embassies. In the discussions **contrasting views about the usefulness of the centralised evaluations and their use came up**. The people had unanimous view about the quality of the centralised evaluations: the evaluations contain plenty of useful information and are considered valuable and well-justified sources of information.

The **perception about their usefulness, however, does not translate into use**. Not everyone who would benefit from them actually uses them. The discussions provided a clear picture: evaluations do not reach the Management Teams or the Offices of the Ministers. This means that amongst the significant non-users are people in positions of power and influence. The active users of the evaluations are experienced professionals of development policy and development cooperation who draw benefit from the reports. Commonly, they work as Director Generals, Deputy Director Generals, Directors of Unit or Senior Advisers at the department and unit level. These are the key Ministry people who define actions, allocate resources and design policies and strategies that would presumably benefit from all evaluations. Yet, discussions revealed that not all of them read the evaluations. The ones who read and use them do so when they are relevant to respondent's work and meet their information needs. They also mentioned that they know how to read them – not full reports, but just summaries, conclusions and recommendations.

> **Evaluations contain useful information but not all people in the Ministry use them.**

Bilateral partner countries, i.e. countries that have a Country Strategy and Country Programme have provided much evidence for the evaluations in the past. Almost every evaluation commissioned by the EVA-11 wants to conduct a country visit and it has become more difficult to identify case study countries for the evaluations. The **people working in the Embassies**, for most part, **do not find the centralised evaluations useful** for their work because the Embassies and in-country partners do not receive concrete recommendations for improving their work. The country strategy evaluations formed an often cited exception; they are considered very relevant to the Embassies.

Views of external stakeholders in Finland

This chapter focuses on the views of MFA's external stakeholders, namely the people from the Development Policy Committee (DPC), Parliament, National Audit Office, Civil Society Organisations and private sector actors.

The external stakeholders of the Ministry have a high degree of appreciation towards the evaluation reports: they recognise that the **reports contain valuable information**. Among the Development Policy Committee members and secretariat the information and knowledge provided by centralised evaluations is regarded important.

However, **majority of the people involved with the Development Policy Committee do not use the evaluations**. The evaluations are on the agenda once a year when the Annual Report that the EVA-11 prepares is discussed. Individual evaluations are not discussed at the DPC meetings. The Working Groups of the DPC are known to use evaluations selectively, i.e. if the evaluations are relevant to the tasks at hand. Members of the Parliament expressed a high interest in the reports but do not use them in their work. A centralised evaluation has been on the agenda of the Foreign Affairs Committee of the Parliament only once during the review period from 2015 to 2022.

Finnish CSOs utilise the evaluations for multiple purposes.

Several **Finnish CSOs** have found ways to **utilise the evaluations for multiple purposes.** They study all reports regardless their scope and topic with the purpose of using the reports for their own organisational learning. They also look for opportunities to utilise the results in their programmes, advocacy work and in the dialogue with the Ministry.

Private sector stakeholders follow-up centralised evaluations occasionally, typically when it is directly linked with their work and business. Those who participate in the DPC receive an update on the evaluations once a year when the Annual Report is presented, and others learn of the evaluations when an invitation to a publishing event arrives from the EVA-11. So far, they have not found uses for these evaluations.

3.1.2 Recommendations of evaluations

The findings in this chapter are based both on documents and stakeholder discussions. The review ToR focuses on the recommendations of the centralised evaluations and the degree to which they are relevant, realistic and actionable.

Main categories of recommendations in the evaluations

The EVA-11 commissioned 23 centralised evaluations during 2015-2022 (see Table 1 in Chapter 2.1). Four of them were meta-evaluations with a distinct focus on evaluation quality and they are not included in the data of Figure 1. Reviews and assessments were not analysed because there are only three reports in the sample.

The **recommendations made in 19 evaluation reports** were structured to nine main categories. Figure 1 below highlights the overall view of the recommendations in these evaluations. The categories with a high number of recommendations (more than 20) are: implementation and aid instruments (how the Finnish development policy and development cooperation is implemented); MFA staff, management and leadership; development of policies and strategies; and influencing and collaboration with stakeholders. Three categories had a moderate number of recommendations (between 10 and 20) and consist of Results-Based Management (RBM), financial resources and regulations and Human Rights Based Approach (HRBA) and Cross-Cutting Objectives (CCOs). Only a couple specific recommendations were made about evaluation practices. Importantly, no recommendations were made about impact, sustainability or relevance in these evaluations (ninth

category). The centralised evaluations' concern is clearly on how the Finnish development policy and development co-operation is implemented, and on the number and competence of people in the MFA and in Finnish Embassies, and the MFA management and leadership.

It should be noted even if a category has many recommendations it does not mean that same recommendations have been repeated. Rather it reflects common issues that are important to address regardless the evaluation topic or aid instrument.

Figure 1. Overview of recommendations provided by 19 centralised evaluations 2015-2022 (total 161)
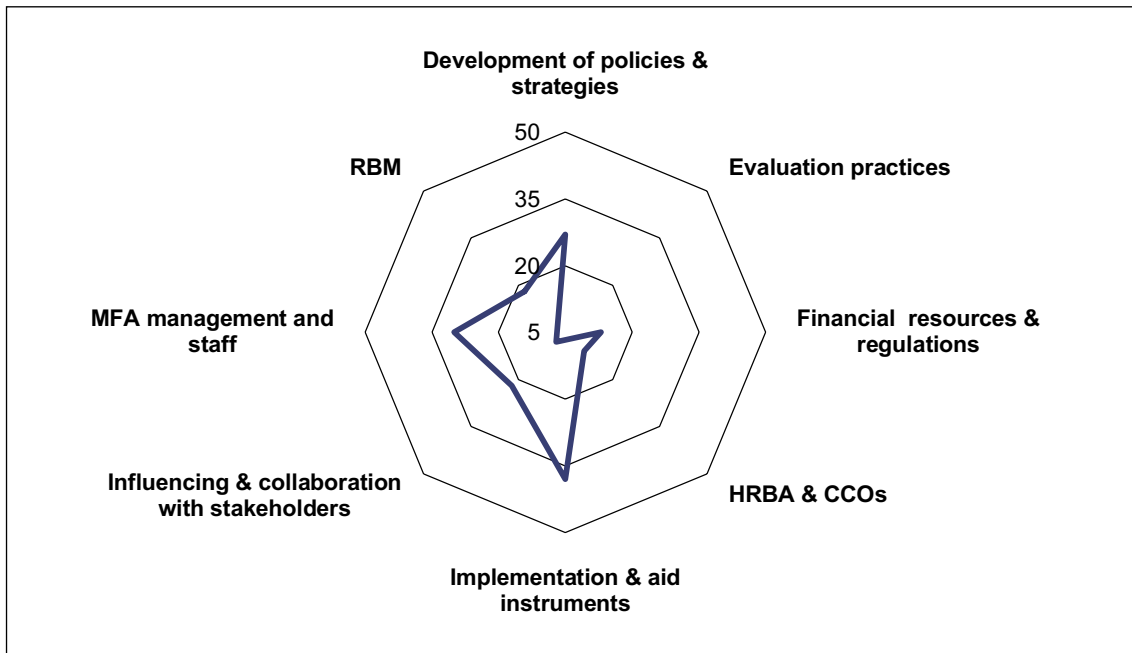
Source: Review team.

The four **meta-evaluations** made a total of 36 recommendations, with a majority **focusing on improving evaluation practices**. Only a handful (between 1 and 5) recommendations were made on sustainability of projects and programmes, HRBA and CCOs, RBM and on finances. This is interesting considering that assessing the quality of Finnish aid has been a significant part of the ToRs of all four meta-evaluations since 2015. The meta-evaluations place more emphasis on assessing the quality of decentralised evaluations than on the quality of Finland's development cooperation.

The number of recommendations has fluctuated from 3 in the developmental evaluation of the Business with Impact Programme (Halme et al. 2019) up to 14 in the water Diplomacy evaluation (de Man et al. 2021). In some older reports also sub-recommendations were made which meant that more than 20 recommendations were made. Several persons considered this number too high. This practice has been checked which reflects the attention the EVA-11 has given to improving the quality of recommendations. On an average, there are 8 to 9 recommendations per report.

In 2019 the EVA-11 introduced **the Findings, Conclusions and Recommendations workshop** into the evaluation process. The purpose is to provide the evaluation teams with an opportunity to receive feedback from the intended users before the reports are drafted (MFA 2021b). The workshop seems to serve its purpose, for example by weeding out recommendations that are not feasible.

Actionability is understood to consist of practicality, feasibility and clarity. The analysis of the reports suggests that the **recommendations have become more focused and clearer** which is expected to make them more actionable. The review of the evaluation reports suggests that the quality of recommendations and style of presenting them has improved over time. The feasibility of the recommendations is further tested in the management response preparation (see chapter 3.1.3).

Issues related to people working in the Ministry and Embassies repeated in the recommendations

The issue of people (in the category MFA staff, management and leadership), having too few people working at the Ministry and/or in the Embassies, was consistently raised in the evaluations. Half of the evaluation reports (12 reports out of 23) bring this issue up in recommendations, either focusing on the people available and/or on their capabilities. In addition, in many reports the issue is strongly referred to in the conclusions but without a corresponding recommendation.

The EVA-11 has analysed evaluation reports completed during 2018-2020 (MFA 2021b). The findings match closely with the findings of this review. The four main categories of repeat recommendations were:

- Steering the planning and implementation of development cooperation and using the knowledge-based tools (theories of change, data-driven analyses of partners or the situation in the partner country and elements supporting adaptive guidance, such as risk mapping and financial flexibility);

- Staff capacity building through training and guidelines;

- Matters related to staff planning, administration and resources; and

- Means and coherence of policy influencing.

Views from the users

According to the MFA people, the evaluations commonly recommend a new strategy paper and address MFA staff related matters. About **the quality of recommendations a mixed message** was received: some people opined that the recommendations have improved while others argued that there is still room and scope for further improvement. The relevance of recommendations is directly linked with the perceived usefulness of the evaluation. If the report and recommendations fit well with the responsibilities and tasks at hand, they are considered very relevant, otherwise not.

According to the EVA-11, the evaluation teams are advised to provide recommendations only to the MFA. This is because the management response process of the MFA does not apply to external stakeholders. It is, however, possible to provide the Ministry with recommendations on influencing external stakeholders. External stakeholders have also participated in the work of management response groups. For example, the Embassy of Finland in Vietnam and Business Finland participated in the preparation of the management response for the transition process evaluation in Vietnam.

3.1.3   From recommendations to agreed actions

This chapter focuses on the management response preparation and the extent to which the actions committed to in the management response adequately address the recommendations. The findings arise mainly from documents.

Management response preparation and contents

The **preparation of the management response** begins after an evaluation report has been approved and published. The Evaluation Norm (MFA 2015a) states that "*in these management responses it is considered to what extent the management of the MFA needs to react to the recommendations*". This purpose is also adopted in the Toolkit for Management Response Working Groups (MFA 2021c). A management response working group should be set after the

> **Management responses focus on addressing evaluation recommendations.**

EVA-11 has approved the evaluation report. The members of the working group are expected to represent all those actors whose inputs will be required to implement the recommendations. The first duty of the working group is "*to jointly decide which recommendations they consider relevant and propose to be approved and which in their opinion should be disregarded / disapproved*." The Toolkit prepared in 2021 also defines that the management response should be completed in three months. During the period of review, the management response has consisted of two documents, a management response decision (published in the MFA website), and a follow-up table with information about agreed actions (available upon request). The draft management response prepared by the group is discussed at the Development Policy Steering Group. The Under-Secretary of State responsible for development policy signs it (MFA 2015a).

9 out of 10 recommendations lead to action

The review analysed 20 management responses that were completed during 2015-2022 (see Annex 5 for a list). Only the follow-up tables were looked at, not the signed management response decisions. This is for two reasons. Firstly, because the follow-up tables can be compared with the evaluation recommendations and with the follow-up reports of management responses. Secondly, the signed decisions do not provide that information. The management response decisions are too vague, and it is difficult, if not impossible to assess what has been agreed to take forward of the recommendations and what not. The signed decisions do not reflect the recommendations of the evaluations, nor do they identify the responsibilities or deadlines for actions.

A total of 175 recommendations were made in 20 evaluation reports. The management response groups agreed fully with 63% of the recommendations (110 recommendations) and those recommendations led into actions to implement them. Another 10 % of recommendation (18 recommendations) are such that action was already taken place (partly or fully). 18 % of the recommendations were either partly agreed or partly rejected (32 recommendations). Finally 9 % of recommendations (15 recommendations) were either clearly rejected or no action planned. This suggests that **about 90 % of the recommendations were somehow considered relevant to the Ministry**. This percentage, however, should be taken with a pinch of salt.

In the sample there are only few examples of justified rejections of recommendations that are clearly spelled out. Even if the management response group does not agree with the recommendations

at all or agrees only with a part of the recommendation, there seems to be a compelling need to design some kind of actions. This then leads into many vague actions that miss the gist of the recommendation. Subsequently, the managers at various levels of the organisation become over-loaded with implementing these actions. Such actions have a minor role in introducing changes in collaboration, policies, instruments, etc.

Recommendations on financial resources and people and their capacities are systematically rejected in the management responses. This issue will be returned to in chapter 3.3.2.

The review of the management responses confirms that **the main function of the management response is to address recommendations.** The process is not looking at the evaluation and the information contained in the report as a whole.

> **The Toolkit for Management Response Working Group offers clear guidance for the members of working group.**

There are also **issues with the follow-up tables.** The writing style is vague, actions are agreed but it is not exactly clear to what extent they address the recommendation. In many follow-up tables responsibilities and deadlines are clearly set, but there are also many where the responsibilities are divided between many units and explicit deadlines are not set. These same issues subsequently influence the follow-up reports (see chapter 3.1.4). The EVA-11 reported these same issues to the Development Policy Steering Group in 2019 (MFA 2019):

- Definition of measures and responsibilities in management responses: Actions are often described in the passive "is clarified, is activated, is encouraged, is confirmed". The expression often leaves open the concrete measures needed to achieve change.

- The responsibility for implementation and reporting remains unclear if several actors are mentioned in the responsibility column.

- Strong commitments to take action are challenging. The responses highlight good intentions, i.e. "Let's try to do".

- In some management responses, the deadline for interlinked actions was the same, although such actions cannot take place at the same time.

The 2021 Toolkit for Management Response Working Groups is considered a major improvement for the process. One implication is that management responses are completed with less time than before. The guidelines are well-prepared and **offer clear guidance for the members of the working group**. The follow-up table is a tool for the responsible MFA units and departments. It is well structured and serves also as a reporting tool. It provides space to group members to record their justifications on why action is taken or not taken, describe the actions and commit to implement the actions. The format has a slot for everything if you know how to use it. In the sample there are some that are perfectly done: first, there is a clear statement about the recommendation – agreed or not – with justifications. It then leads into clear, practical actions for which one department or unit takes responsibility for within an agreed deadline.

3.1.4 Implementation of agreed actions

The final step in the evaluation process is the tracking of management responses and reporting. The analysis looks at the extent to which the actions committed to in the management response are implemented. The findings are mainly based on documents.

Follow-up reports and the guidance

The EVA-11 guidelines for management response process also provide instructions for follow-up reporting (MFA 2021c). The reporting period has been two years. Based on the Evaluation Norm the responsibility to follow the implementation of the management responses remains with the EVA-11. The Unit also uses the follow-up reports as a basis for reporting to the Development Policy Steering Group.

Outcomes of 12 completed management response processes have been looked at using the follow-up reports as evidence (see Annex 5 for a list). Analysis of follow-up reports suggest that **virtually all actions agreed in the management response have been implemented**. In these 12 evaluations, 111 recommendations were made. 106 of them were partly or fully agreed to when the management response was prepared (96 %). According to the follow-up reports, a vast majority of actions agreed have been implemented by the responsible units. Actions for 92 recommendations have been implemented as planned and actions for 13 recommendations have been partly implemented. This leaves only one recommendation with a not implemented status. This was in the management response for the Meta-evaluation for project and programme evaluations 2012-2014. A planned action about organising a coordination session for advisers in the Ministry about evaluations and evaluation capacity was not organised.

> **Actions agreed in the management response have been implemented, but their follow-up reports do not tell about the intended change.**

The numbers look good, but they tell only a part of the story. The **reports do not tell what the intended change was that the actions should have produced**. The element of change is missing from the follow-up reports as they do not contain any analysis about the effect of implemented actions.

On the positive side, for the Ministry people one of the benefits of this process is that **it forces the management of relevant units to take stand**. The experience of the people was that if an evaluation identifies a drastic need for change, then action will be taken, even without a management response decision.

The findings of the review match perfectly with the EVA-11 discussion paper (MFA 2021b) where evaluation reports and management responses from 2018 to 2020 were studied. The paper argues that:

- The management response matrices attached to the management response decisions are used in the follow-up. As a tool it keeps the recommendations alive. However, the summaries done in the matrices are mechanical.

- The practice reduces the burden on reporting for the units, but there is no or little space for joint analysis and learning. The report is only discussed at the Development Policy Steering Group.

- Monitoring is done late, i.e. at the earliest two years after the management response has been developed. The risk of institutional memory disappearing is evident.

- The practice of follow-up reporting works in terms of accountability. The benefits and effectiveness of monitoring are limited as there is no or little analysis on the observations and good practices that were highlighted in the evaluation reports.

- The management responses and follow-up reports play a very small role in bringing about or accelerating change. The most important thing seems to be that the need for change remains in the discussion agenda.

This review also agrees with the issues that the EVA-11 brought to the discussion in the Development Policy Steering Group in 2018 (MFA 2018d):

- Instead of the present style of reporting actions, a **report on the impact of the evaluation and the lessons learned** afterwards could be considered. It would be good to think about how the follow-up reporting of evaluations could be developed to inspire more learning.

- The template for follow-up has not always made monitoring easy. One reason is that **the actions defined in the follow-up table are often detached from the major strategic decisions and changes**. Management decisions that clearly state the responsible party and the timetable are easier to follow.

- The purposes of the follow-up reports could be reconsidered: **does it serve mainly accountability or is it another opportunity to learn from the evaluation**? Revisiting the management response two years after its approval would be an opportunity to reflect on the findings and recommendations raised by the evaluation, and how they have helped to strengthen strategic thinking and planning for the evaluated activity.

Commitment to implement actions

The discussions did not provide any evidence on how the commitment to implement the actions manifests itself at best. Analysis of documents provides only limited evidence. The follow-up reports indicate that respective departments and units have a high commitment to implement the agreed actions. Among the good examples are the country strategy and RBM evaluations: their recommendations have fed into important ongoing processes, that have led into concrete action plans, and the plans have been implemented.

3.2   Factors affecting the uptake of evaluations

This chapter looks at the **Review Question 3: What are the typical factors facilitating or hampering the uptake of centralised evaluations?** Analysis of the evaluation process is also incorporated in this chapter (part of RQ 2). The sub-questions focus on the factors, on reasons for not taking action on recommendations in the management response, and the way actions agreed in the management response are implemented are addressed in this chapter.

Key findings

The EVA-11 has put a lot of emphasis on the improvement of evaluation process. This is reflected in the improvement of tools, guidelines and training provided to the expected users of evaluations.

Ideas for evaluation are identified in a participatory manner. The final selection of the evaluation topics is in the hands of the EVA-11.

Earlier only large evaluations with a long time to implement were commissioned. People in the Ministry support the new practice of also including reviews into the portfolio.

The outsourced Evaluation Management Service (EMS) has evolved into conducting evaluations in a standardised manner and even quality.

Open communication and opportunities to participate either as informants or reference group members in evaluations was appreciated by many. Also the Findings, Conclusions and Recommendations workshops were considered useful.

Evaluation function and the potential evaluations provide for learning and knowledge-based management is not well understood in all the departments of the Ministry.

The reports contain plenty of useful information that can be utilised for many years. Evaluation users expect improvements in style of reports, access to evaluations, and receiving tailor-made information.

Ownership issues are evident in the management response preparation. After the evaluation report has been approved, changing the ownership from the EVA-11 to the department/unit leading the management response process has not always worked well.

Half of the evaluations made recommendations about the people or their capabilities and many others made findings or conclusions on them. The MFA leadership has not acted on the issue.

The management response process works well in addressing recommendations that clearly fall within the mandate of Development Policy Steering Group and departments and units engaged with development cooperation. It does not work for recommendations dealing with the issue of people of the MFA and/or the Embassies.

3.2.1 Evaluation planning and management

The findings presented in this chapter arise from the discussions and focus on distinct phases of the evaluation process. Many different factors, some enabling and some hampering, play a role in the use of evaluation results. Contradictory views that came up in the discussions have been reflected here as well.

Planning of evaluations and supportive tools

The EVA-11 has introduced and developed **many evaluation tools and guidelines and provided training** on evaluation. The goal is to facilitate learning during the evaluation process and to support all evaluations and evaluation capacity in general. Examples of tools are listed in Box 2. These tools indicate that the EVA-11 has a solid commitment to continuous improvement of evaluation methods and process contributing to improvement of evaluations, their use and usefulness.

Box 2. Examples of tools developed for different phases of the evaluation process

Evaluation manual. It is web-based combining documents, graphic and videos, and serves as a training package (MFA 2023).

eOppiva basic evaluation training course (MFA 2023)

Guidelines for Evaluation Reference Group (MFA 2020b)

Findings, Conclusions and Recommendations workshops to discuss tentative findings, conclusions and recommendations before the evaluation report is finalised (also applied in this review)

AHA-tool for collecting information on evaluation results by Priority Areas of Finnish Development Policy (MFA 2020g)

Checklist for recommendations for better evaluation recommendations (MFA 2023)

Toolkit for management response working groups (MFA 2021c)

Practical tips for addressing cross-cutting objectives in evaluations: The Implementation of the Human Rights Based Approach; and Gender Equality and Non-Discrimination (MFA 2023)

Practical tips for addressing Cross-Cutting Objectives in evaluations: Climate Resilience, Low Emission Development, Environment and Biodiversity (MFA 2023)

Source: Review team

Based on the review of Annual Evaluation Plans 2015-2022 and on discussions it is evident that the EVA-11 has consistently applied **a participatory process in preparing the Annual Evaluation Plan and in selecting the centralised evaluation topics.** The Unit has systematically expanded the list of internal and external stakeholders that are consulted. For the Annual Evaluations plans prepared during 2015-2018, the EVA-11 consulted relevant departments and units in the Ministry

(MFA 2015b, MFA 2016a, MFA 2018c). Currently the EVA-11 also seeks the views of the Offices of the Ministers, Development Policy Committee, the umbrella organisation of Finnish civil society organisations Fingo and National Audit Office (NAO).

> **EVA-11 applies a participatory process in preparing the Annual Evaluation Plan.**

The **final selection of topics of centralised evaluations** is in the hands of the EVA-11. When the Unit prepares the annual plan, it considers the potential of different departments and units to participate in the evaluations, its own capacity to manage the evaluations and the budget it has available for evaluations. In the past plans all or almost all financial resources have been earmarked for selected evaluations already at the beginning of the calendar year. Thus the EVA-11 has had limited chances to respond to any urgent information requests, arising, for example, from the MFA leadership.

During the review period, on an average three to four evaluations have been completed per calendar year. Concerns were raised that this number is high compared to **absorption capacity of the Ministry**, the capacity of the organisation and people to benefit from the evaluations. Such a steady flow of evaluations puts demands on the time of the people whose inputs are needed during and after the evaluations as informants, members of reference groups and members of management response groups.

The evaluations are also not the only information and knowledge sources relevant to the organisation. Other departments and units produce analyses and commission assessments and research papers, auditors come up with their reports, and the National Audit Office produces management audit reports, etc. In general, **information overflow is a problem for all**. Working together with other units and actors was proposed by some MFA people: could the evaluations be part of a more cohesive knowledge product steam? To avoid duplication of efforts the EVA-11 has two-way consultations and shares work streams with the Development Policy Committee and the Unit for Planning and Research. These relationships could be developed further for evaluations to feed into their work and vice versa.

> **Disruptive change calls for a wide range of evaluative analysis.**

Timely availability of evaluation results was frequently mentioned as a factor that would contribute to improved use of evaluation results. The public sector, and the MFA is no exception, faces the same rapid and disruptive changes that are already experienced in the private sector. An elaborate planning process used to be a strength and advantage in the past, but it is now becoming a liability. At present it takes approximately two years from the identification of the evaluation topic before the results are finally available and can be shared with expected users. In the meantime, the urgency and demand for the results may have passed.

Starting from 2021 the EVA-11 has also conducted **reviews that provide results quickly**. The reviews have a narrow focus and short time frame compared to the evaluations. This review is an example: the duration was just four months from start (October 2022) to results presented in a form of draft report (February 2023). The other three are the Business-to-Government Review (Makkonen 2021), the review of Finland's cooperation supporting peace processes (MFA 2021f) and the assessment of the Response of Finnish Development Policy and Cooperation to the COVID-19 Pandemic (Laaksonen et al. 2022). Ideas for them have come from the Under Secretaries of State and/or from the Director General of Department for Development Policy.

This **practice of conducting reviews was welcomed by many**. The EVA-11 could balance the portfolio of its assignments between the long and in-depth evaluations and the rapid reviews. Several MFA people suggested that the EVA-11 could indeed produce very useful reviews for the benefit of not just the development policy practitioners but also the Offices of the Ministers and Management Teams of the Ministry. The 2023-2026 Evaluation Plan contains an example of such a review that is expected to feed into pertinent information needs (the review on development policy as part of foreign and security policy).

Evaluation Management Services arrangement

Only the persons who work as coordinators, evaluation team leaders or evaluation managers in the EVA-11 offered views about the outsourced Evaluation Management Services arrangement.

When introduced in 2017 it was something new. The EMS arrangement has multiple purposes (MFA 2020a): it is expected to increase quality and coordination of the evaluations that the EVA-11 outsources. It is also expected to increase flexibility, efficiency and effectiveness in planning and commissioning the evaluations. The EVA-11 team was satisfied with the arrangement. They also considered that the EMS has contributed to improved evaluation quality.

Information bias is very much evident in the way the evaluations have been conducted: much emphasis is given on data collection using multiple methods, interviewing large numbers of people and studying many documents. In comparison the evaluation teams have allocated less time for data analysis and having e.g. joint writing workshops. Addressing these issues is in the hands of the EVA-11 itself as the Evaluation Manager. The review team considers that such problems can be addressed already in the Concept Note and ToR of an evaluation and during the inception when evaluation teams prepare their work plans and methodologies.

The view of the Service Coordinators and Team Leaders was that the EMS concept provides a tried and tested process for conducting evaluations in a standardised manner. In their experience, the emphasis, so far, has been on producing evaluation reports. The inclusion of Emerging Evaluator posts in the teams is most welcome: it gives Finnish young professionals a chance to build experience in these evaluations.

For evaluation assignments managed through the EMS very senior international consultants have been sought. It is the view of the review team that when the future of EMS is planned the qualifications for the evaluation consultants are worth considering. Requirements for outstanding international substance and/or technical experience could be better balanced with requirements for knowledge and understanding of the MFA, the Finnish Embassies and their functions. Both are necessary prerequisites for high evaluation quality.

Conducting evaluations

People appreciate **the opportunities to participate and engage in the evaluations**. In the Ministry, many mentioned that the evaluation interviews and workshops organised to collect data were useful to them. They considered the dialogue with the evaluators and the questions asked useful because it gives them opportunities to think and reflect. Thus, the evaluations interviews are learning opportunities for individuals.

Many people appreciated the **evaluation reference groups**. The groups were considered a good forum for dialogue, both during the preparation of the evaluation and its implementation. However, there is a risk that some people may become overwhelmed because they end up being part of many different reference and management response groups.

Open communication during an evaluation increases ownership amongst the expected users. It was suggested that perhaps the evaluators could schedule regular briefing sessions with the potential users. This was considered a measure that could also increase ownership of the evaluations among the expected users. Such sessions would also provide timely access to emerging information among the users.

The stakeholders consider that the quality of evaluations has improved in the reporting period. Potential users find evaluations relevant because they produce **impartial information that has been analysed by external experts**. This adds to the credibility of the evaluations.

The **Findings, Conclusions and Recommendations -workshop** engages Ministry people in the dialogue about evaluation results and their potential utility already when they are being processed. The FCR-workshop came up in some discussions, and indeed the Ministry people consider it useful.

The present writing instructions that the EVA-11 has produced for the centralised evaluations set the expected length of the reports at 60-80 pages. The average length of the main reports in the sample (23 reports) is 75 pages. The **current reporting style**, although informative, **does not support use of evaluations**. Although evaluations are not research, the reports are written in a scientific study style. A consensus among the internal and external stakeholders prevailed: the reports are too long. They also found the practice of presenting detailed methodologies at the beginning of the report distracting. The methodologies may be of interest for evaluation professionals but not for the intended users. A focus on concise / shorter reporting would respect the time of the reader.

Using and sharing evaluation results

During the review it was argued that the Ministry has a strong evaluation culture. In fact, in the departments and units working with development policy and development cooperation, evaluation function is widely accepted and considered useful. This applies also partly to the Department for International Trade. These departments and units are the ones that have a strong tradition and culture. But **the evaluation function is not equally well understood in other policy areas and departments**. In the departments working with foreign policy and security policy, evaluation function is not practiced. It was implied that often evaluation is seen as controlling, like inspection and as such a threat and not to be practiced.

The **evaluation process does not move seamlessly from the completion of the evaluation to utilising the evaluation results**. During the evaluation, the EVA-11 has the ownership of the evaluation as the Evaluation Manager. But the EVA-11 takes a step backwards when the management response process starts, and the ownership and leadership is expected from the department or unit leading the process. Persons who have served in the reference groups do not always continue as members of the management response group although they would have the best insights into the evaluation results. Otherwise no clear opinions were expressed about the functions of the

management response groups in the discussions. Issues with management response process will be returned to in chapter 3.2.2.

A cornerstone for the use of the evaluation is that the reports are publicly available. It was not common knowledge that all the reports, together with the related communication products (presentations and briefs) and the management response decisions are available in the MFA website. Yet, this has been the practice of the EVA-11 for many years, and the Unit has also communicated about this in different MFA forums and public events. The stakeholders argued that **the evaluation results should be more easily available**, as at present the potential users do not know from where exactly to find the reports. The view of people in the Ministry and in the Embassies was that the reports are not easily found from the internal systems they are expected to use. External stakeholders receive information about the evaluations if they are invited to the launch event. OpenAid was suggested as an alternative site for providing public access to the reports and knowledge products.

For **communication about the evaluation results** new practices are called for. This was voiced by all stakeholders both inside and outside the Ministry. Ideas that came from the discussions include advance information on coming publishing of an evaluation report, especially to the Offices of the Ministers but also to other potential users, tailoring information according to different groups, and producing different types of short info messages in various media channels. The examples of such short messages ranged from 3-5 bullet points (e.g. for the Parliamentarians) to a 1-pager for the Offices of the Ministers. Overall, there was a huge demand for short but repeated messages about the evaluations across different communication channels.

> **New practices are called for sharing evaluation results.**

It also became clear from the discussions that **the shelf life of an evaluation report is too short.** The reports contain important information that is useful and can be used over a number of years, not just when the management response is prepared. The present processes of utilising information provided in the evaluations direct attention to recommendations and their implementation. There is too little attention on evaluation findings. This is also the direction that the table format used for presenting the evaluation results (findings, conclusions and recommendations) guides the potential users to look for, particularly if involved in preparing the management response.

Meta-evaluation of decentralised evaluations 2017-2020

A lot of commonalities exist between the findings of this review and the findings of the latest meta-evaluation that looked at the usefulness and timing of decentralised evaluations. According to Väth et al. (2022) three main factors facilitate the usefulness of those evaluations: the evaluation design (nature of the evaluation and Terms of Reference), management response process and its follow up, and competence of the evaluation team. The usefulness of evaluation reports often remains at an individual level. Exploitation of evaluation results for organisational learning inside MFA is hampered by a lack of a functioning institutional knowledge management system.

Ideas for future evaluations

Many forward-looking ideas came up in the discussions. **A strong call was made about addressing policy coherence more systematically in the strategic evaluations**. Development policy is part of Finland's foreign and security policy; thus all these policy segments should build on the same values and be human rights based. Yet, it is only development cooperation and humanitarian

assistance that have been evaluated. The stakeholders felt that plenty of potential exist for expanding the scope of centralised evaluation to the other policy sectors and this would feed into improved coherence as well. The challenge at present is that in the evaluation norm development evaluation is tied to ODA-eligible funds.

Policy coherence could be addressed more systematically in the evaluations.

The evaluation professionals saw some other needs for changes, e.g., **a variety of assessments and tools** that serve the rapidly changing aid instruments and contexts could be considered. On case-by-case basis rapid assessments would also work.

All stakeholders indicated a demand for more focused and lighter **reviews** and would welcome reader-friendly quick analyses. There is also interest in on-going **developmental evaluations** (future orientation, feeding forward) and in **synthesis reports** of MFA's own evaluations. Little is known about, for example, about the European Development Fund (EDF) or other EU cooperation, as evaluations of these are not readily available. It was suggested that also synthesis reports of other financiers' reports could be produced.

The people felt that that the evaluations have addressed accountability well. They argued that it would be important to look beyond accountability and **focus increasingly on learning** in the future.

3.2.2   Uptake of recommendations

This chapter builds on the analysis of management response preparation and action taken on management responses that was discussed in chapters 3.1.3 and 3.1.4.

Reasons for not acting on recommendations in the management response

The management response working groups have not agreed with two types of recommendations, namely recommendations made on financial resources and recommendations made on people issues.

In the post-2015 situation it is no surprise that **recommendations for increased funds** for an instrument or a programme or starting a new programme / project have not been taken on board. Such recommendations have all been rejected in the management responses. These recommendations have dealt with different aid instruments or aid modalities. For example, the Evaluation on Improvement of Women's and Girls rights (Rassmann et al. 2018) would have wanted the MFA to consider dedicated medium-term funding for CSOs and Evaluation of the Country Strategy Modality in Fragile Situations (Betts et al. 2020) recommended increased financial flexibility for fragile contexts. Adequate justifications were provided for not agreeing with them in the management responses.

The real issue is **the uptake of the repeated recommendations made about people issues**. They are a hot potato in the management response process. If the recommendation is about capabilities, a standard approach has been to plan actions for staff training, for development of new tools or for network. If the recommendation is about having more people for the job, the management response groups have adopted different tactics in side-stepping this issue. For example, if the evaluators have made a convoluted recommendation that contains two or three parts and one of them deals with the MFA or Embassy people, the people part is ignored, while for the other

> **People issues are systematically overlooked in the management responses.**

parts actions are designed. Recommendations focusing purely on the number of people available for important tasks at hand may be entirely ignored (the silent treatment) or they are addressed with quasi actions that miss the point.

There are examples of management responses where the actions related to people issues have been directed to the Administrative Services, but here lies the problem: the group preparing the management response does not have the mandate to make any Ministry-level decisions about human resources. This applies also to the Development Policy Steering Group and the Under-Secretary of State responsible for development policy and cooperation that are involved in the finalisation and signing of the decision. This is also confirmed in the analysis of the EVA-11 (MFA 2021b). A typical recommendation that transgresses departmental decision making powers is related to increasing number of people, (re)allocation, or rotation. Implementation of such recommendations would require measures other than those purely related to development policy or the implementation of development cooperation.

There are **other reasons for not acting.** The management response process is extra work that is not necessarily prioritised by Director Generals of Departments or Directors of Units. This came up in the discussions. The practice was condoned with given the context of having too few people for the tasks at hand.

Also, there is an ownership issue: the role of the EVA-11 is to serve as a secretary of the management response group, but the Unit is also in charge of initiating the process of establishing the group. On centralised evaluations, the EVA-11 may have the strongest ownership of the management response although the Department/Unit leading the group should assume it. Some good examples do exist of demand-driven management response processes where the unit in charge of implementing the recommendations owns the process (e.g. the recent preparation of the management response for the humanitarian aid evaluation).

Reasons for not implementing actions that were agreed in the management response

As was discussed in chapter 3.1.4 almost all actions that are agreed in the management response are also implemented.

In the discussions with the Ministry and the Embassy people some problems with the management response process came up. The two-year period designed for implementing the management response is too long. Rotation and change of people interrupts development work. Monitoring and follow-up was considered problematic in general and filling up and updating the tables is not in anyone's interest. Management responses are done but it is not sure what came of it (what change?). Even if there is interest for change, then it is just another paper. The fragile states evaluation came up as such an example: the evaluation said that existing instruments do not work well in fragile states, yet no action for change has been taken. For the process to work better, clear scheduling would be needed and the management responses together with justifications for actions should be widely shared.

The views shared in the discussions focused only on the existing process. No ideas for improving the process came up.

3.3   Transformative capability of evaluations

This chapter looks at the **Review Question 1: How transformative are centralised evaluations?**
It looks at two dimensions, the extent that the evaluations have influenced/and or changed policies,
guidelines and processes in the MFA, and the extent that the evaluations have influenced and/or
changed policy dialogue inside and outside the MFA.

Key Findings

The evaluations have influenced development cooperation policies, guidelines, and pro-
cesses in the MFA. Clear, practical themes of evaluations have been relevant to operational
units and departments.

Timing of evaluations was raised often. The best examples of evaluations that have had
plenty of influence in the MFA are those that have fed into ongoing decision making or im-
plementation processes.

The meta-evaluations have fed into development of evaluation manual, tools and training.
Otherwise their influence is limited.

The 2021 Report on Development Policy Across Parliamentary Terms and its parliamentary
preparation process is an example of a successful policy influence inside and outside the
MFA. Otherwise the centralised evaluations have only modest influence on policy dialogue
inside and outside the MFA.

The Management Group or Extended Management Group have not discussed the results
of the evaluations as a knowledge-based management practice. The evaluations are ad-
dressed at the level of departments and units engaging in development policy and devel-
opment cooperation.

The EVA-11 informs the Offices of the Ministers about completed evaluations, but the eval-
uations are not discussed with them. Same applies to the Development Policy Committee
and Parliament and its Foreign Affairs Committee.

Coherence between foreign policy, security policy, trade policy and development policy has
been addressed in a couple of evaluations. For evaluations to be able to contribute to the
improved policy coherence at the Ministry, also foreign, security and trade policies together
with development policy should be addressed.

3.3.1 Influence on development cooperation policies, guidelines and processes

The people working at the MFA and in the Embassies cited **many fitting examples of evaluations that have had influence or have contributed to changes in specific policies, guidelines or internal processes of the MFA**. Their views are supported by documentary evidence as well. Among them are the 2018 and 2022 development policy result reports to the Parliament: both reports make ample reference to findings and recommendations of evaluations (MFA 2018b, MFA 2022b). Similarly, in the 2021 Report on Development Policy Across Parliamentary Terms (MFA 2021a) evaluations are frequently referred to. It was also learned that evaluations and their results are used, for example, in the work of the Quality Group (MFA 2022c), in country programming (MFA 2020c), and in influencing multilateral organisations (MFA 2020d). Quite a few people also mentioned the reform process of operational modalities in development cooperation (commonly known by its Finnish acronym *KeTTU*) in this context.

According to the people, the best examples are the evaluations that have fed into the development of Results-Based Management approach and tools at the MFA (the results-based management evaluation in 2015 and the knowledge management evaluation in 2019). The country strategy evaluations (Kenya in 2015, country strategies evaluation in 2016 and fragile contexts evaluation in 2020) have influenced development of subsequent country strategies and programmes. Business Finland considered that the developmental evaluation of the BEAM programme in 2019 gave them useful inputs for implementation of the programme. Finnfund had similar experiences about their evaluation in 2018. Also, the evaluation that addressed development cooperation carried out by the Department of Russia, Eastern Europe and Central Asia (2021) has helped to improve cooperation and processes within the Department.

The follow-up reports of evaluation management responses (13 reports available, see Annex 5 for a list) were analysed to **identify further uses of evaluations** (Table 2). Different types of guidelines stand out as the most common ways of use (country programming, CSO development cooperation, Quality Assurance Group, Cross-Cutting Objectives, triple nexus and humanitarian aid). Evaluations have contributed to the Results-Based Management, risk management and Human Rights Based approaches. Evaluations influenced both the content and the parliamentary approval process of the Report on Development Policy Across Parliamentary Terms in 2021. Evaluation recommendations have also fed into the implementation of the KeTTU process. In light of the management responses, among the influential evaluations are the country programme evaluations (Kenya, Fragile states and Country strategies), the 2015 RBM evaluation, Women and girls evaluation in 2015, Meta-evaluation in 2017, Forced displacement evaluation in 2019 and the series of CSO evaluations that was completed in 2017.

Table 2. Examples of evaluation use in the MFA

EXAMPLES OF USE REFERRED TO IN THE FOLLOW-UP REPORTS	NUMBER OF MANAGEMENT RESPONSES	NAME AND YEAR OF EVALUATIONS
Country programming (guidelines, planning, reporting)	6	Fragile states (2020), Women and girls (2018), Country strategies (2016), Inclusive education (2015), RBM (2015), Kenya (2015)
CSO development cooperation (instrument description, application criteria, reporting guidelines)	4	Fragile states (2020), CSO (2018), Inclusive education (2015), RBM (2015)
Reform process of operational modalities in development cooperation (*KeTTU*)	4	Women and girls (2018), Country strategies (2016), Meta (2017), RBM (2015)
Results-Based Management (Guidelines, Theories of Change, Bilateral project manual)	3	Fragile states (2020), Women and girls (2018), RBM (2015)
Development policy (contents, parliamentary approval process)	3	Fragile states (2020), Forced displacement (2019), RBM (2015)
Quality Assurance Group guidelines	3	CSO (2017), Meta (2017), RBM (2015)
Triple nexus guideline	2	Fragile states (2020), Forced displacement (2019)
Guidelines for Cross-Cutting Objectives	2	Women and girls (2018), CSO (2017)
Human Rights Based Approach	2	Meta (2015), CSO (2017)
Humanitarian aid guidelines (incl. humanitarian mine action)	2	CSO (2017), Humanitarian mine action (2015)
Risk management	1	Fragile states (2020)
Financing mechanisms	1	Fragile states (2020)
Multilateral influencing plan	1	Women and girls (2018)

Source: Review team

Timeliness of evaluations and how evaluations should feed into decision making processes was often raised by the MFA and Embassy people. The best influence that any evaluation can have is when its results feed into an ongoing decision making or implementation processes. Timing is crucial - if evaluation is too early, the information may be forgotten. It is definitely too late if the evaluation comes after the decisions are made. The potential to contribute information and knowledge to the decisions is lost. The CSOs seconded this view.

Not so good examples

Although the evaluation on management of results information and knowledge (Palenberg et al. 2019) has made an impression and has remained in the memory, it could have been more influential. The MFA has faced difficulties in finding mutual understanding on how to proceed with the recommendations as is indicated by the one and half year delay in completing the management response decision. Another example, where **all intended users did not have timely access to evaluation results** is the evaluation of the CSO programmes that receive multiannual programme-based support from the MFA. The evaluation was commissioned in three separate lots and the last lot was completed too late. The reports were not available to all CSOs when needed for new programme planning.

Meta-evaluations with limited influence

Four meta-evaluations have been conducted during 2015-2022. Their scope has varied: the objectives of the first three meta-evaluations (Cooney et al. 2015, LeBlanc et al. 2016 and Silvestrini et al. 2018) focused on assessing the quality of decentralised evaluations and the related planning documents, and the quality of Finland's development cooperation. The lat-

> **Meta-evaluations need a new focus.**

est meta-evaluation (Väth et al. 2022) assessed also actual use and usefulness of decentralised evaluation reports. Only few people brought up meta-evaluations in the discussions: their opinion was that **the meta-evaluations no longer have much value, and they are repeating the same issues report after report**.

The meta-evaluations have fed into the evaluation manual (latest update from 2018) and into evaluation tools, guidelines and training. This is evident from the document analysis but in the discussions these issues were not raised.

Decentralised evaluations are those commissioned by the other departments and units of the Ministry and the Finnish Embassies. In last two meta-evaluations (Silvestrini et al. 2018 and Väth et al. 2022) less than 50% of the evaluations were commissioned by MFA. In the latest meta-evaluation 53 % of the reports that were assessed for evaluation quality were commissioned by multilateral international organisations under the UN, CSOs (local and international NGOs), the World Bank, intergovernmental agencies and development cooperation agencies from other countries. When the report is commissioned by others, the role of the MFA is only in commenting the ToRs and the evaluation reports. The MFA has in fact no real role in quality of the report. It is the view of the review team that if the practice of commissioning meta-evaluations is continued, it could be utilised for other purposes. Demands for different types of synthesis type of analyses came up in various discussions. Meta-evaluations could serve that purpose very well.

3.3.2 Influence on policy dialogue

The findings in this chapter are based on the discussions. Building on discussions with the EVA-11 during the inception, the review adopted a broad understanding of policy dialogue. The question was considered not just from the angle how the evaluations have influenced implementation of development policy within the MFA, but how the evaluations have influenced the formulation of development policy both inside and outside the MFA.

Influence inside the MFA

Many views about policy dialogue came up in the discussions with people working at the MFA, but only one concrete example of centralised evaluations directly contributing to the dialogue was cited. The preparations of the Report on Development Policy Across Parliamentary Terms (MFA 2021a) was mentioned as policy dialogue process to which centralised evaluations have contributed. This is an excellent example, but unfortunately the only one. The people working in the MFA did not offer any other similar examples. They did make comments about indirect ways of influencing the dialogue, i.e. the evaluations contain valuable information, they strengthen the knowledge base, and this is considered valuable.

Quite a few Ministry people raised **the coherence between development policy, foreign policy, security policy and/or trade policy** and mentioned that it is yet to be addressed in the evaluations. They fully supported the statement that development policy is an integral part of Finland's foreign and security policy. They also argued that there is a reciprocal relationship: the other policy segments contribute to development policy and should be addressed in a more coherent manner there. For the evaluations to be able to contribute to improved policy coherence in the Ministry, also **foreign, security and trade policy aspects should be addressed in the evaluations**. They also reminded that results-based management is at the core of development policy and cooperation and could serve as a model for foreign and security policy as well.

The PATH2030 - evaluation of Finland's sustainable development policy (Berg et al. 2019) that was commissioned by the Prime Minister's Office was cited as a good example of evaluation addressing policy coherence at a national level. The contributions that foreign policy, including development policy, have made to Finland's Agenda 2030 are analysed there.

Coherence issues between development policy and foreign policy or trade policy have been addressed in some recent evaluations. Among them are the Evaluation of Economic Development, Job Creation and Livelihoods (Laaksonen et al. 2021), the review of cooperation supporting peace processes (MFA 2021f) and the Evaluation on Water Diplomacy (de Man et al. 2021). According to the EVA-11, the latest Annual Evaluation Plan (2023-2026) also contains one review that will address the coherence between development, foreign and security policy.

The **evaluations** managed by the EVA-11 **are defined as "comprehensive and strategic"** (MFA 2015a). The term "strategic" was interpreted in many different ways in the discussions. Some people from the MFA considered these evaluations strategic because they appear on the agenda of the Development Policy Steering Group. Others thought that because the evaluations are of interest to the management of the Department for Development Policy, that makes them "strategic", and that the evaluations therefore inevitably contribute to policy dialogue.

The EVA-11 has a practice of informing the Offices of the Ministers when an evaluation has been completed and the results are made public. However, there is no practice of discussing them further. It was also learned that the individual evaluations and strategic issues identified in them are not brought to the attention of the Management Team or Extended Management Team of the Ministry. The discussions on evaluation findings, conclusions and recommendations remain at the level of departments and units that are in charge of development policy and development cooperation.

Influence outside the MFA

There was a consensus among the Development Policy Committee, the Secretariat and a number of members, namely Parliamentarians, CSOs and other interest groups, that **the evaluations do not contribute to the dialogue about development policy**. Individual evaluations are not discussed at the Development Policy Committee meetings. Thus, the DPC has not worked as a channel to communicate evaluations results to the Foreign Affairs Committee of the Parliament or the Parliament in general. The working groups of the DPC are known to use the reports on case-by-case. Thus, the reports can have some influence within the working groups. The discussions clearly indicate that there is room for working together: the Development Policy Committee consists of a wide range of stakeholders that are interested in the evaluations and would be interested in spreading the message further.

Evaluations do not feed into development policy dialogue in Finland.

Evaluations have indirect influence on the policy dialogue with the Parliament. Evaluations are frequently referred to in the 2018 and 2022 development policy result reports and the 2021 Report on Development Policy Across Parliamentary Terms that were discussed in the Parliament.

The EVA-11 informs the Offices of the Ministers (Minister for Foreign Affairs and Minister for Development Cooperation and Foreign Trade) about completed evaluations. However, the evaluations are not necessarily discussed with them. Parliamentary elections are due in April 2023. A concrete suggestion was made for the meetings and discussions about the contents of the new Cabinet programme: the EVA-11 could utilise the meetings as a platform for feeding information and knowledge obtained from the evaluations into the process.

Ideas for future evaluations

Several examples of where evaluative evidence would be useful to substantiate dialogue and decision making were provided by the DPC members. Among them are the thematic priorities of Finnish aid, the aid instruments and how their respective balance has changed during the period of the review, such as the shrinking role of bilateral development cooperation and emergence of new instruments for private sector engagement.

3.4   Knowledge and learning from evaluations

The fourth and final review question is: **How could the MFA better capture and retain knowledge and learning from evaluations as one tool for knowledge-based management?** The findings are based on discussions. In the stakeholder discussions, the people were encouraged to look forward and share their views and ideas on the topic.

Key findings

The evaluations have a lot of potential for supporting knowledge-based decisions at the level of Ministry leaders, but the potential is not reached because the Management Teams and Offices of the Ministers do not use the evaluations. The evaluations have benefited knowledge-based decision making at the department and unit level.

Individual efforts to generate knowledge and learning from the evaluations prevail over collective learning solutions in the Ministry.

Retaining of knowledge and learning is influenced negatively by information overflow, difficulties in finding the evaluations from the MFA information systems and website and contextual factors of the Ministry (rotation of post holders, continuous change of special public servants and understaffing). There is lack of time for learning, reflection and retaining knowledge and reflection. The evaluations are not able to transfer knowledge they generate across boundaries in the organisation.

Participating in evaluation reference groups can be a learning journey at best: inspiring ideas and collaboration make participants to implement changes already during the evaluation. The working groups preparing the management responses do not provoke similar insights or learning.

So far only a few units and departments embrace the idea for using evaluations for learning, connecting learning to their context and using it to make changes.

There is plenty of potential and ideas for capturing knowledge and learning from the evaluations. The EVA-11 team has a vision of centralised evaluations as an essential part of knowledge management in the Ministry. The EVA-11 team has the potential to influence the decision makers: it has access to all relevant decision making, coordination and advisory groups.

The EVA-11 has already experience from collective learning solutions that can be used for nurturing learning and using knowledge from evaluations.

3.4.1 Learning from evaluations

Frameworks of learning, organisational risks, and new co-creation models that inspire the review team thinking are introduced in Annex 7. Those frameworks have been applied for analysis in this chapter.

Learning is the foundation for change and has positive impact on knowledge-based decisions.

Learning is an important way of utilising evaluation results, and learning is regarded as the foundation for change and having positive impact on knowledge-based decisions. The role of evaluations for learning and knowledge-based decision making is prominently raised in the recent OECD Review of Evaluation Systems (OECD 2023). This role has also been recognised in some earlier evaluation and reviews as well. According to Palenberg et al. (2019), the MFA leadership were generally supportive of strengthening evidence-informed

learning and decision making but only a few "champions" considered and pushed for it as a priority for the MFA as organisation. However, the "leadership" refers not the leadership of the Ministry but to the leadership in development policy and coordination and administration (the Department for Development Policy, Department for Africa and the Middle East and Department for the Americas and Asia and the Administrative Services). According to the Peer Review of Four Evaluation Function (2020) utility viewpoint insists that evaluation needs to be as seen as a continuous process from issue identification to institutional and individual learning. Also, the Finnish government level aim is to increase knowledge-based management and decision making as part of modern leadership (Ministry of Finance 2023).

The function of the development evaluations in the MFA is manifold: the evaluations are expected to serve accountability, learning, improvement of quality, and transparency and openness. Importantly, **development evaluation is expected to serve the organisation-wide learning from development cooperation experiences** (MFA 2015a). Both the discussions and the document review support the finding that accountability need remains important, but emphasis of learning from evaluation is becoming increasingly important. This is reflected in the way how the EVA-11 has organised the evaluation process (see chapter 3.2).

The evaluations provide a lot of data and information for the Ministry, and to stakeholders in Finland and abroad. It is clear to all that generating knowledge and learning from evaluation results takes time and practice. But to many, **"learning" translates to "organising training"**. Learning is then understood as "taking in information", i.e. listening to a lecture or reading a text. Both the discussions and the review of management responses suggest this.

> **Many challenges for learning from evaluations exist.**

Individual and collective learning should interconnect to achieve a sustained change in the working culture of the MFA. At present **personal mastery is evident** with an abundance of examples from individual efforts to generate knowledge and learning from the evaluations. This is particularly the case with the people whose background is in the development cooperation, but it is not necessarily the same with career rotating diplomats. They did not identify the evaluations as important sources of information for their work, possibly because in functions for foreign policy and trade policy evaluations are not practiced.

The **learning from evaluations is challenged by many factors in the MFA**. Information overflow came repeatedly up in the discussions. Evaluation reports are not the only important knowledge products that are competing for attention. Evaluations are interesting only if they provide relevant to me now -results. The evaluation reports are not easily found from the information systems. The contextual factors of the Ministry, that is, rotation of post holders, continuous change of special civil servants, understaffing, emphasis on informal discussion and relying on email as learning tools, were also amply referred to in the discussions. Both learning and knowledge-based management suffer as a result of these factors.

Team learning activities take place ideally in various working groups and events in the MFA. Participating in reference groups of evaluations can be a learning journey at best. Inspiring ideas and collaboration make participants to implement changes already during the evaluation. Working groups preparing the management response have not been mentioned at all in this context; there was no talk about those groups as provoking insights and learning, although the EVA-11 has advocated for that in the Development Policy Steering Group. The concern of same people

participating in reference groups and working groups for management response was also raised as this is considered extra work and delegation of duties is not done because there is not enough staff.

Effective adult learning calls for **self-reflection and reflection of learning**. Most people talked about their workload and hectic workdays and how it hinders use of evaluation results. Time for learning and reflection practice is very limited and not valued in the Ministry. Some people implied that individually the solution is to leave the MFA.

Very little about **the role of evaluations contributing to shared vision and systems thinking** elements came up in the discussions. By definition, the centralised evaluations with strategic emphasis are expected to have an impact on development policy, development cooperation, and on how the MFA as an accountable and responsible agency is operating. Pre-conditions for successful change are co-creating shared vision and increasing mutual understanding of systemic impacts. For that to happen, time and space for learning and knowledge generating is needed at the team, unit, department, and ministry level. So far, a few units and departments embrace the idea of using evaluations for learning, but the rest of the organisation has not yet quite grasped it.

3.4.2 Retaining and using knowledge from evaluations

Several types of problems that hamper both knowledge-based decision making and utilising information from centralised evaluations came up in the discussions. The challenges that people in the Ministry have experienced in benefiting from the evaluation results are presented in **the framework on learning friendly culture** in Table 3 (Senge et al. 1999). Ten distinct challenges or sets of forces oppose use and utility of evaluations. As a matter of fact, all these challenges are predictable. Although they often appear as seemingly independent events, they are interconnected and interdependent. It is also argued that progress is sustainable only if innovators and change agents learn to understand why the system is pushing back, and how people's attitudes and perceptions contribute to the pushback.

> **Systemic pushbacks for learning culture are evident in the MFA.**

Table 3. Ten interdependent challenges opposing change in the MFA

CHALLENGE APPLIED ALONG SENGE ET AL. (1999)		EXPERIENCES OF THE MINISTRY PEOPLE
Initiating change Factors preventing use and utilising of evaluations, and learning	1. Control over one's time	The challenge of work time kept popping up in every discussion: *I don't have time for this.* For example, no time to meet with the review team, no time to participate in reference group, in meetings, in evaluation publishing event, in reading evaluation report, etc. It was an exception if time issue was not raised in the discussions.
	2. Inadequate coaching, guidance and support	Inadequate support was repeatedly mentioned. *We need help but…* Centralised evaluations are not easily found in the MFA information systems. The reports are too long and difficult to grasp because of jargon and style. Also, people are expecting someone else to tailor the findings for them.
	3. Relevance	Perceived lack of relevance was among the main challenges for not using and utilising evaluation results. Evaluations are used *only if it is relevant to me and directly to my work right now.*
	4. Management clarity and coherence	This challenge deals with understanding the evaluations within broader systems in the MFA, the MFA leadership and administration, especially the role of Administrative Services that was raised in discussions. Utilising evaluation results for knowledge-based decisions at the leadership level of the Ministry is missing.
Sustaining the change Sustaining the use and utilising of evaluations	5. Fear and anxiety	Many causes of fear were raised: what future parliament election will bring about, how the units manage the more demanding tasks and increasing responsibilities, or fear of speaking up. Challenge of anxiety in evaluations came up, e.g. between the evaluation team and the MFA people the viewpoints can be many: *This is not quality evaluation if …* or *The evaluators don't understand us at all.*
	6. Negative assessment of progress	Centralised evaluations face the challenge of negative assessment among the expected users. *For whom these are made? This is too theoretical. This is not working for our needs.*
	7. Isolation and arrogance	Evaluation teams and the EVA-11 face this challenge as evaluation managers and owners of evaluations. The challenge extends itself into the management response process. *We have the right way. They don't understand us.*
Rethinking and redesigning These are about learning from evaluations and making change accordingly	8. Prevailing governance structure	Maintaining the existing structure is seen important and evaluation teams are advised not to rock the boat. Alternatively, the independence of the evaluation teams is challenged by asking them to be the messenger. *Who is in charge of this stuff? There are prohibiting acts and laws… but please bring it up.*
	9. Diffusion	The diffusion challenge translates to evaluations not being able to transfer knowledge generated across boundaries in the organisation. At the highest, the evaluation results reach the department level only.
	10. Organisational strategy and purpose	Occasionally evaluations have managed to win this challenge, e.g. in influencing on how the MFA is reporting to the Parliament and wider public or in the adoption of Results-Based Management.

Source: Review team

Problems and risks associated with sub-optimal use of evaluations in learning and decision making

In the discussions with people working in the Ministry and the Embassies many views on how the MFA operates support the fact that the evaluations are not feeding into decision making at the highest level. Depending on the position of the stakeholder different types of risk were emphasized, but all these issues are **risks for the MFA as organisation**. The three categories of risks identified by Mintzberg (2018) are used in the analysis, as follows:

> **Organisational risks increase if evaluations are not used in learning and decision making.**

1. *There is no strategy at all, and work is fragmented*: The concept of "leadership" was blurred among the MFA people that talked to the review team. When they mentioned the "leadership", they were asked to clarify who they were talking about. A variety of responses were received, from unit level up to the ministers. When activities are led at the department or unit level, **work becomes fragmented.** The Development Policy Steering Group has a significant role in the Ministry in coordinating development policy, development cooperation and related coherence issues. Director Generals oversee policy coherence and financial results in their respective geographical and/or policy sector. The issue that remains is how coherence with development policy is addressed in the foreign, security or trade policy activities and how development evaluation could contribute to that.

2. *The strategy is lost if you give up the valid and useful strategy and start working with new things because they are new and interesting*: This type of risk is also evident in the operations of the Ministry. People spoke of many changes in the development cooperation and aid instruments that have taken place without a true dialogue. Important decisions have been made without using the evidence available in the evaluations or commissioning forward looking evaluations on the subject matter. The EVA-11 has commissioned one evaluation with a forward looking angle, namely the evaluation on water diplomacy, Water as an Entry Point for Peace Mediation (de Man et al. 2021), that focused on the opportunities and identifying the ambitions and future interests.

3. *The strategy is wrong, when the organisation is drifting through small steps and ostensibly small decisions into unfavourable position*: Many evaluations have expressed their concern about the issue of understaffing in the MFA, and this review makes no exception. People are constantly overloaded with work because the Ministry has not acted on the issue. Thus the Ministry is **drifting into unfavourable position**. Collaboration between departments suffers when they are made to compete of people and people suffer from work stress.

Stakeholder ideas for future

Many insightful ideas and solutions for nurturing learning from evaluations and knowledge-based decision making came up in the discussions as follows:

- The MFA would benefit from adopting collective learning solutions. These could cover different types of activities and involve stakeholders inside and outside MFA, many of which the EVA-11 has experiences from. Establishing an Evaluation Community of Practice and involving stakeholders also from outside the MFA is among them. Also,

the EVA-11 can engage CSOs and public and private sector partners in the dialogue about systematically using and benefiting from evaluations. Organising morning coffees to discuss evaluations topics will build on dialogue inside the MFA.

- Build a partnership with the Development Policy Committee. It has the potential to be the forum for strategic discussion and policy dialogue with a range of stakeholders.

- Use the momentum that 2023 provides to foster learning and knowledge-based decision making. After the parliamentary elections in April 2023, a new government will be stepping in, with two or three new ministers and Offices full of energy and eager to make difference in the MFA. This is a great opportunity to feed in knowledge from centralised evaluations and about the benefits of the evaluation function and create a positive imprint of the evaluations in the minds of the decision makers.

- Increased flexibility and diversity of types of evaluations would strengthen the link between evaluations and utilising information from evaluations in decision making.

4 Conclusions

Strengths and weaknesses of the evaluation process and evaluations

During the period of the review, the EVA-11 has worked towards improving implementation of evaluations by increasing the engagement of stakeholders, by developing formats, guidelines and tools and by developing knowledge products and training materials. The evaluations and the evaluation process have many strengths that are summarised in Box 3. They reflect the aim of the EVA-11 to develop the evaluations to the direction that they would be utilised more, both for learning and knowledge-based management. The review supports these aims fully.

Box 3. Strengths of the evaluations and the evaluation process

The EVA-11 prepares the Annual Evaluation Plans in a consultative manner and also engages stakeholders outside the MFA.

The types of evaluations commissioned by the EVA-11 have become more diverse, e.g. reviews have been adopted for producing evaluative information.

Attention to coherence between development policy and foreign policy and/or trade policy has improved.

The evaluation process provides opportunities for the MFA staff to participate in the evaluations.

The outsourced Evaluation Management Service arrangement produces standard quality.

The evaluations are conducted by external experts. The reports contain a lot of information and are of high quality.

Source: Review team

The findings of the review indicate that there are also several areas needing improvement. Some minor ideas for improvement were raised already in the respective sections in Chapter 3. Box 4 summarises the priority areas for improvement. The use and utility of the evaluation and how transformative they ultimately will be is only partly in the hands of the EVA-11. The Development Evaluation Unit can, for example, define the type of evaluations that are planned for the year, make sure that results are available on time to feed ongoing decision making processes, and define if the purpose of an evaluation is to feed into accountability, learning or both. The utility of evaluations, however, is ultimately in the hands of the people who are expected to embrace and own the results and apply them in learning and knowledge-based management. In this respect, leadership services both from the political and public servant leaders are necessary.

Box 4. Weaknesses in the evaluations and the evaluation process

The evaluations commissioned by the EVA-11 take approximately two years from the idea to results being available. Evaluation results have not always been available when needed.

Sharing of evaluation results is not responding to a variety of stakeholder needs.

The reports are not appealing to potential users. Interested users of evaluations do not know where to look for evaluation reports and knowledge products.

Past evaluations are forgotten too quickly even though they contain valuable information.

Because the evaluations are not taken up in the Management Teams or by the Offices of the Ministers, the MFA leadership does not generate knowledge from evaluations.

A challenge in transferring the ownership of the evaluations from the EVA-11 to the departments and units responsible for using the evaluations is evident.

Management response process is used for uptake of evaluation results, but the process is not working well.

Evaluation use and utility has focused on recommendations of evaluations with less emphasis on findings and conclusions.

The practitioners of development policy and development cooperation demonstrate a strong evaluation culture. Overall the evaluation culture is not yet fully developed in the Ministry.

Evaluation function is not well understood, and evaluations are not practiced in the departments working with foreign and security policy.

Focus on solely individual training and learning is not effective in fostering knowledge-based management in the Ministry.

Source: Review Team

Sub-optimal influence on policies and policy dialogue

It is encouraging that people both in the Ministry and among external stakeholders consider the evaluation reports and information contents useful, also from knowledge-management point of view. Many evaluations have influenced policies, guidelines and processes of development cooperation. Among the best examples are evaluations that have fed information into ongoing decision making or other development processes (e.g. country strategy development and Results-Based Management processes).

However, the review considers that the evaluations have not completely filled their potential as strategic evaluations. The influence the evaluations have had on policy dialogue is limited. This is because examples about a specific evaluation or evaluations used in the Ministry to influence dialogue about major choices and shifts in development policy were not forthcoming.

Similarly, the evaluations do not influence development policy debates at the Parliament. They are not discussed in the Foreign Affairs Committee of the Parliament or in the Parliament in general. Within the Development Policy Committee the use is limited to the working group level although the information value of evaluations as such is recognised and supported. The Development Policy Committee and the Ministers can provide the mechanism to bring the relevant information from evaluations to the political debates.

In the discussions with the Ministry people, the "leadership" was understood to consist of Director Generals of departments and Directors of units but seldom reaching up to the top. The departments and units that are involved with development policy and development cooperation evaluations discuss and utilise evaluations, but the use does not extend beyond the Development Policy Steering Group. The review considers a major shortcoming that the evaluations are not discussed at the Management Teams of the Ministry or with the Offices of the Ministers.

It is encouraging, that the EVA-11 has a seat in all relevant groups, on those working on management of the Ministry (Extended Management Team), improving quality of development cooperation (Quality Group) and coherence of development policy (Development Policy Steering Group). With the profound knowledge of past and planned centralised evaluations EVA-11 people have potential to influence as informal leaders. To be fair, they are already doing this to some extent and with success. The team collaboration works smoothly in the Unit and the team shares a vision of centralised evaluations as an essential element contributing to the MFA knowledge management. This is very helpful for increasing the transformative power of evaluations in the future and was also evident in the process of drafting the Theory of Change for the Unit (see Annex 2).

The Finnish foreign policy consists of many policies that all go hand in hand, including development policy, security policy and trade policy. It is welcome that the centralised evaluations have already started addressing coherence between different policies. If continued it is likely to contribute to improved policy dialogue as well.

It is for the reasons described above that the evaluations cannot be regarded as transformative. The politicians and the Ministry have not used the evaluations when directions of Finnish aid, development cooperation aid modalities or aid instruments have been discussed and agreed upon.

Management response process is flawed

The review acknowledges that the EVA-11 has facilitated the management response process with appropriate guidelines. The improved process, however, has several problems. It is not fit to address the recommendations that transgress the institutional boundaries of the Ministry. It focuses mostly on recommendations paying less attention to findings. All centralised evaluation results are processed in the same way without considering how operational or strategic the recommendations are. The process contributes to accountability, but it does not instigate learning or knowledge-based management and evidence of change produced through the management response is limited. The two-year time period allocated for the process is too long to keep focus on activities. It remains open if the process as such is useful for all centralised evaluations. Evaluation results mutually approved e.g. in a reference group during the evaluation work lead into changes at any case and recommendations are implemented despite the management response not being signed when mutual understanding is reached.

The limited reach is the biggest shortcoming in the current process. Major institutional, strategic or policy coherence issues cannot be fully addressed by the management response process that

only reaches the level of departments. The review also expects that in the future evaluations would focus on wider issues across the Ministry. Thus, the evaluations need to be discussed in respective forums, i.e. also in the Management Teams.

Information sharing and knowledge products

The level of evaluation use came as a surprise to the review team because it is at a lower level than what was expected. In the Ministry the active users of the evaluations in the department and unit level consist of experienced practitioners of development policy and cooperation. Also the people of the stakeholder organisations use the evaluations but to a varying degree. For example, some Finnish CSOs use centralised evaluations in a professional way and take the useful information and knowledge for improving their own work or advocacy activities.

Use of evaluations would be facilitated if the reports are easily available. The review was astonished to learn that many internal and external stakeholders were not aware that the reports are publicly available in the MFA website. Evaluation use would be also facilitated if the results are summarized well, and if the reports are available in a timely manner. At present the use is hampered by the volume and writing and reporting style. The use as well as implementation of recommendations is also hampered by absorption capacity in the Ministry because too many evaluations have been ongoing per year in past several years.

Stakeholders acknowledge evaluation reports as source of information and knowledge. Although the EVA-11 expects the intended evaluation owners to take over as soon as the report gets published, the EVA-11 still has the strongest ownership on the reports. This is challenging because at the end of the day the EVA-11 is responsible for disseminating development evaluation results but is not alone responsible for facilitating learning from evaluations. For example, the Unit for Development Policy is responsible for improving knowledge and competence related to development policy and development cooperation and coordinating related training in the Ministry. Evidently, there is a need to find mutual understanding on how to move on together with learning approach.

The Development Evaluation Unit has a seat in the right tables and has all the power to improve sharing of evaluation results inside and outside MFA. Bringing people on board to share their ideas and comments publicly takes time and practice. What remains inside the Ministry and is open to outside stakeholders are design questions. The higher in decision making the stakeholder is, the higher is the importance in sharing stakeholder-tailored information in advance, with easily perceived form, and systematically. In arranging public events, planning them from participant viewpoint is important - what is the added value this event provides to them? The EVA-11 could benefit from making an influencing plan that would include known decision making points, would identify the key stakeholders, why they are important for evaluation use and utility, and what evaluation results are available for them. The EVA-11 could also start producing quick, frequently circulated Readers Digest type of briefs of present and past evaluations for wider audiences. This would help to tackle the information overload.

Evaluations could have a longer shelf life. The EVA-11 can lengthen the life of an evaluation report by raising interesting topics to discussion forums, by arranging morning coffee events with thematic agenda and bringing information from various past evaluation reports into discussion. Another channel is to remind new people of the existing evaluations and where to find them. Perhaps evaluation info could be part of briefing materials (and now we don't mean the reports as such, only few lines introductory text what the evaluation is about). The Development Evaluation Unit can be information hub for the evaluations commissioned by actively promoting evaluation results from new evaluations but also promoting results from previous evaluations as well when situation

arises. Synthesis of all new analyses (evaluations, studies, reviews, etc.) could be commissioned and published every six months to keep these information and knowledge alive.

Knowledge and learning from evaluations

At present the use of evaluations focuses on implementing recommendations via a management response. This overlooks the evaluation report as a source of information and knowledge that would be very relevant for both knowledge-based decision making and learning. The role of the MFA leadership is crucial in fostering more learning-oriented culture and knowledge-based management. Leadership is like sailing: you can only succeed when you know two things. You need to know where you are right now, and you need to know where you are aiming at. Evaluations provide answers to the first question and ideas about the second one.

In many evaluations, important institutional issues came up. From the viewpoint of how the Ministry as an organisation is operating, the role of the MFA leadership is of crucial importance in providing leadership services. Knowledge-based leadership services are needed to ensure that the whole organisation is operating for the wellbeing of people in the MFA and that they can provide their full potential to their duties and tasks,that ultimately make the Finnish foreign and security policy and development policy relevant, impactful, and sustainable.

Improving knowledge-based management and learning from evaluations in the Ministry calls for an improvement in learning approach: instead of providing training and information for individuals a collective learning approach needs to be adopted. Team and group-based dialogue and learning has wider institutional effects. Team learning takes place in various working groups and events.

Arranging learning processes can start with a question on how to use the existing material. In adult learning, the most effective methods are team and group based sharing, experimenting in practice and reflecting. Ask more why? questions instead of what? and how? Learning needs to be supported and encouraged by the leaders. Instead of independent single events learning processes that last about four to six months should be developed. Three elements are fundamental in learning: it needs to be a process, it needs to be arranged in teams or groups, and it needs to be facilitated. Learning process can use various models, address various topics, and be virtual or face-to-face.

When strengthening the learning culture of MFA, learning objectives should be integrated with the MFA performance management procedures. Learning objectives should be set at individual, team and unit level. The focus of the MFA leadership is needed on this kind of system level development.

Space for flexibility in the plan and diversity of evaluation types

There is a mismatch between the speed of disruptive change and the pace of current evaluation process. Evaluations will be used if they are timely available and addressing to direct information needs. Increased flexibility in scale, focus and time is needed. The move towards diversified evaluations is already taking place in the EVA-11. Perhaps the next step can be also to diversify reporting and moving away from "pure" reports towards other types of knowledge products. The future evaluations can consist of different types of reviews, meta-analyses and synthesis reports, developmental evaluations, self-evaluations, peer reviews, citizen boards, they will be all needed to name a few.

Improvements in reporting style would be needed, i.e. shifting from technical style of reports to more reader-friendly reports and other types of information and knowledge products. It will also support the EVA-11 in producing tailored information to stakeholders.

5   Recommendations

The recommendations focus on evaluation use and utility and on activities that the EVA-11 can initiate within the coming six months and can keep up continuously. As was argued in the conclusions, the EVA-11 needs action and support from the MFA leadership and from other departments and units as well. There are six recommendations under four areas.

Improved utilisation of evaluations for knowledge-based decisions

1. Engage the new Ministers and people in their Offices and raise their interest in the significance of the centralised evaluations.

Parliamentary elections and the new Cabinet starting its work in spring 2023 is the perfect opportunity for the EVA-11 to brief both the political and public servant leaders of the Ministry about the value of evaluations in providing information for knowledge-based decision making.

2. Constantly share evaluation information and knowledge with the Offices of the Ministers, Management Teams, and in the Development Policy Steering Group.

The EVA-11 can encourage the MFA leadership in the Offices of the Ministers and the Management Teams to use evaluation results in knowledge-based decisions and knowledge management. All evaluations have findings that are useful to the Ministry, but it is particularly important that the leadership is informed of evaluation results when they deal with MFA wide institutional issues, policy coherence issues, and when leadership services are needed to address an issue.

3. Initiate a process of rethinking and revising the management response process.

The limited reach of the management response is the biggest problem. Major institutional, strategic or policy coherence issues cannot be addressed in a process that only reaches the department level. It is also expected that future evaluations will address wider issues that need to be discussed in respective forums, the Management Teams among them.

Improved sharing of evaluation results

4. Improve on the information sharing about the evaluations to respond the varied needs of different stakeholders

Revisit the communication and information sharing practices of EVA-11 both during the evaluations and after the evaluations are finalised. Engage the Department of Communications in the process. Use external support for developing communication tools, materials and sharing methods.

Feed the MFA leadership with advance information of every evaluation and ensure that points are easily absorbed. Share actively with the Development Policy Committee and other stakeholders inside and outside MFA. Also keep reminding where the information is stored and updated.

Share evaluation results in such a way that you raise interest and attract your stakeholders to look for more information. Provide material that is focused, brief and easily perceived. Share the key messages, just 3-5 bullet points would be appreciated. This needs continuous attention so keep on repeating. Use both direct channels like e-mails, and increasingly institutional channels like the MFA webpage, intranet, digital discussion platforms, and social media. Collect also feedback on the methods and materials and revise accordingly.

Learning from centralised evaluations

5. Prioritise collective learning oriented methods in engaging MFA stakeholders

Learning from centralised evaluations calls for learning oriented methods and practices. The MFA has already at hand EvalSpace platform that can be used for virtual collaboration and learning. The EVA-11 can, for example, organise a series of facilitated group discussions, create a Community of Practice in EvalSpace, and organise morning coffees to feed evaluation results forward for discussion collectively and systematically. External stakeholders could also be invited in such discussions. It would be appropriate to draw attention to all evaluation results, not just to recommendations or the latest evaluations.

The dialogue can take many models, so it is important to start experimenting facilitated dialogue, and then reflecting and modifying it.

Co-creation with other teams, units and departments is needed because the EVA-11 is not alone responsible for this. The information sharing practices brought up in recommendation number 4 also contribute to learning.

Space for flexibility and diversity to meet the information needs

6. Increase flexibility and diversity of evaluations

Evaluations will be used more if they are available to address direct information needs and are timely available when needed. The annual evaluation planning has an important role in providing predictability both to the EVA-11 work and for stakeholders, but at present there is a mismatch between the fast speed of disruptive change and the slow pace of current evaluation process. For the next Annual Evaluation Plan, pilot a practice of leaving some resources unallocated initially. Engage with stakeholders and the MFA leadership and promote evaluation function and reviews than can be completed in a few months' time.

The move towards diversity of evaluations and reviews is welcome and should be continued. In the future, the EVA-11 could produce different types of reviews, synthesis reports, developmental evaluations, meta-analyses, self-evaluations, peer reviews, and use citizen boards in the process, to name a few.

Annexes

Annex 1. The Review Team

Team Leader / Senior Evaluator: Ms Kristiina Mikkola with an extensive experience from development policy and cooperation evaluation, planning and implementation, also from key external stakeholder groups. As Team Leader, Ms Mikkola was responsible for the management of the team and organisation of the work. She was also the lead writer of all reports and deliverables.

Senior Evaluator / team member: Ms Päivi Äijälä has development policy and cooperation experience added with profound experience in co-creating individual and institutional change, change management and organisation development. Ms Äijälä took the lead in the design and leading of the co-creation process and contributed to all other activities, including reporting.

Quality Assurance Expert: Ms Jaana Vormisto (FIANT Consulting Oy) with extensive experience from development policy, Cross-Cutting Objectives and Human Rights Based Approach, planning and evaluations, organisational change, utilisation of research and evaluation results for effecting change. The QA Expert reviewed the draft inception report and draft and draft final reports of the review.

Annex 2. Terms of Reference

23.8.2022

Terms of Reference (ToR)

Review of the Use and Utility of Centralised Evaluations

1. Introduction

EVA-11 (EVA-11) of the Ministry for Foreign Affairs carries out centralised strategic level and policy evaluations, meta-evaluations as well as syntheses of decentralised and centralised evaluations. In addition, different types of preliminary evaluability or background reviews on wider evaluations, or on other issues of importance to the organisation, are commissioned by EVA-11.

The use of evaluations is high on the agenda of EVA-11, prompting the commissioning of this review of the use and utility of centralised evaluations. The review was also one of the actions included in the management response of the evaluation of knowledge management[1].The review thus complements the analysis of the use, utility as drivers of change and tools for knowledge management.

The ultimate **purpose of the review** is to strengthen the use of evaluation evidence for decision making, learning and accountability and to enhance the use and utility of centralised evaluations. Ultimately, the review contributes to improvement of the implementation and effectiveness of development policy and cooperation.

2. Context: Evaluation function and evaluation process

The development evaluation norm (2015) guides evaluation function and evaluations in the Ministry for Foreign Affairs (MFA). The norm stipulates that evaluations serve both organizational learning and accountability. The norm also guides the preparation of annual plans for development evaluation as well as evaluation processes, including management response and follow-up of management response. Further guidance on the evaluation process is given in the development evaluation manual (updated in 2022).

EVA-11 prepares **annually a comprehensive plan of evaluation** in cooperation with other departments and units conducting development evaluation. EVA-11 also consults other stakeholders, e.g. Development Policy Committee. The annual plan includes a plan for centralised and decentralised evaluations. EVA-11 brings the comprehensive evaluation plan and possible subsequent changes to the Development Policy Steering Group for discussion and presents it to the Under-Secretary of State for confirmation.

The MFA commissions all evaluations and reviews to independent evaluators and expert teams. For centralised evaluation the EVA-11 has a Framework contract for **Evaluation Management Services (EMS).** The rationale for the EMS is to provide a service-based management and facilitation of evaluations. The role of the MFA is a commissioner of evaluations (planning the evaluation and leading the design and implementation of evaluation, quality assurance of evaluation

---

1   MFA. "How do we learn, Manage, and Make Decisions in Finland's Development Policy and Cooperation" Management of results information and knowledge at the Ministry for Foreign Affairs (2019)

deliverables, dissemination of results and recommendations of evaluations and follow-up of the use of the evaluations). The role of the EMS Service Provider is that of a manager (administration, coordination, facilitation of evaluation teams and missions, quality assurance of the evaluations as well as Help Desk and backstopping services) and implementer of evaluations.[2]

For each centralised evaluation, a reference group is established. The overall **purpose of the evaluation reference group is to support a useful, credible, transparent, impartial and quality evaluation process**. Establishing a reference group is a good way to engage and involve internal and external stakeholders representing a range of expertise and/or end users of evaluations. Having a reference group accompany the evaluation process means that stakeholders are encouraged to participate and bring their perspectives, expertise and concerns to the table in a collaborative spirit.[3] Another way to engage the users of evaluations and enhance the learning and use of evaluations has been to organize Findings, Conclusions and Recommendations workshops to discuss tentative recommendations before the evaluation report is finalized.

All development evaluations carried out in the MFA are responded to by a management response. EVA-11 facilitates the process as its secretariat. When a centralised evaluation is ready the EVA-11 requests that the department responsible for the activities consults comprehensively other stakeholders and establishes a working group and nominates the chair for the group. Based on the recommendations of the evaluation, the group prepares a draft management response, which takes into account the roles and responsibilities of all actors for the future changes. The chair of the working group takes the draft management response to the Development Policy Steering Group for comments after which the response is finalised. The chair of the working group presents the management response to the Under-Secretary of State for approval.

The activities agreed in the management response are followed up for two years after which the responsible department prepares a report of the implementation of the activities. EVA-11 has developed a toolkit for evaluation managers and management response working group[4].

All centralised evaluation reports are published on the MFA's website and stored in the electronic archive (AHA-KYT). AHA also includes information on evaluation results by Priority Areas of Finnish Development Policy.

3. Results of previous evaluations and reviews

The joint peer review of the evaluation units of Switzerland, Ireland and the MFA was carried out in 2020[5]. The peer review concluded that evaluation is accepted as important contributor for ensuring projects and programmes are meeting their objectives and proving beneficial in terms of development work. However, evaluations are less understood and supported among staff working with foreign policy and trade. Furthermore, the evaluation work was perceived to be useful and was used by operational staff and management, though this was mostly at the policy rather than operational level. Reaching broader audiences is also challenging. Country teams saw decentralised evaluations as more directly relevant to their work. The peer review found that centralised evaluations are geared toward the broader issues facing the organization than any one set of

---

2   MFA. Terms of Reference: Framework Agreement for Development Evaluation Services (2020)

3   MFA. Guidelines for evaluation reference group (2020)

4   MFA. Toolkit for management response working groups (2021)

5   Peer Review of Four Evaluation Functions (2020)

operations which seem a reasonable focus. The peer review recommended to enhance the planning phase and management response process and improve communication and dissemination of evaluation results.

Furthermore, the peer review noted that evaluation is ultimately expected to be used to achieve practical results through institutional learning and implementation of recommendations. Thus, it is important that the evaluation reports are not only independent and credible, but that they address issues important to MFA decision-makers in a timely way and with clear messages. The peer review found some issues in this regard relating to scope, complexity and duration which drive cost but also have implications for the utility of evaluation results. It also recommended that EVA-11 should exercise more control over the evaluations and diversify its portfolio of evaluations products.

The **evaluation on management of results information and knowledge**[6] found that centralised evaluations, their synthesis in annual reports and performance audits produce useful results information and have successfully influenced decision making at the MFA but do not fully utilise the comprehensive body of knowledge they generate.

Regarding the use of results information for learning and decision making, the evaluation evidenced a moderate level of influence of results information on both learning and decision making at the MFA. Centralised evaluations have had significant influence on decision making at the MFA. However, they are targeted primarily at senior MFA management and the Finnish government and parliament, and their relevance and usefulness to general MFA staff has remained limited. The evaluation noted that the relevance and usefulness could be improved by "filtering" and "packaging" relevant findings and conclusions into suitable formats for different, more specifically defined audiences in (and beyond) the MFA. Furthermore, learning during the entire evaluation process can be further strengthened by increasing staff participation to increase understanding and ownership of evaluation findings and the feasibility of recommendations.

The evaluation found that project mid-term reviews produce useful results information and influence operations if they are conducted well. It also found that meta-evaluations of projects and programmes have become truly analytic and useful. Meta-evaluations provide useful information about, for example the degree and quality with which some of Finland's development policy priorities are covered and derive some good practices and lessons learned regarding institutional and development performance. Also the meta-evaluation of individual CSO evaluations was considered valuable by the MFA's Unit for Civil Society.

The **meta-evaluation of MFA's project and programme evaluations in 2017–2020**[7] further assessed the use and utility of decentralised evaluations. The evaluation revealed that decentralised evaluations are often used for learning in teams, decision making, planning of follow-up, and management of ongoing interventions. The learning stays however more on individual level. Furthermore, the evaluation noted that a large majority of the evaluations is assessed as satisfactory or very useful. Timing of the evaluations, relevance of recommendations, the extent of recommendations being realistic, the learnings from the evaluation, and the implementation of recommendations are positively assessed.

---

6   MFA. Evaluation. How do we learn, Manage, and Make Decisions in Finland's Development Policy and Cooperation. Management of results information and knowledge at the Ministry for Foreign Affairs of Finland (2019)

7   MFA. Meta-evaluation of project and programme evaluations in 2017–2020 (2022)

4. Purpose, objectives and users of the review

The ultimate **purpose of the review** is to strengthen the use of evaluation evidence for decision making, learning and accountability and to enhance the use and utility of centralised evaluations. Ultimately, the review contributes to improvement of the implementation and effectiveness of development policy and cooperation.

The **objectives** of the review are to 1) identify main factors that enable or hinder the uptake of evaluation results and implementation of recommendations both in the MFA as well as among other stakeholders in Finland (e.g. Development Policy Committee, Parliament, civil society organisations (CSOs), 2) identify key themes/issues of strategic relevance to the MFA, for which there seems to be important gaps in implementation of recommendations, and 3) provide concise conclusions and recommendations that help the MFA to step up the use of evaluations and implementation of recommendations.

The primary internal stakeholder is the Development Policy Steering Group which, when exercising its support and coordination function to improve the effectiveness of development policy and cooperation, will benefit from the identification of thematic issues of strategic relevance that require further action. Furthermore, the primary users of the review are the EVA-11 and the senior management of the MFA and secondarily departments managing development policy and cooperation. The review is also of interest to the Development Policy Committee.

5. Scope

The review will cover the period of 2015–2022. In order to assess the use and utility the review will include annual evaluation plans, evaluation reports, management responses and reporting back on the management response, annual reports of development evaluation as well as processes related to them. The review will cover centralised evaluations commissioned by EVA-11 (in total 23 evaluations and 3 assessments, see Annex 1) excluding joint evaluations.

6. Key review questions

The review will address the following questions:

RQ1: How transformative are centralised evaluations?

1.1 To what extent centralised evaluations have influenced and/or changed policy dialogue inside and outside the MFA?

1.2 To what extent centralised evaluations have influenced and/or changed policies, guidelines and processes in the MFA, including with regard to crosscutting objectives of Finnish development policy?

RQ2: How are centralised evaluations used and how useful they are?

2.1 To what extent the evaluation processes (from planning until the follow up of the management response) support use of evaluation results and implementation of recommendations?

2.2 To what extent are the recommendations relevant, realistic and actionable?

2.3 To what extent do the actions committed to in the management response adequately address the recommendations?

2.4 To what extent is there commitment to implement the recommended actions and how does that manifest itself at best?

2.5 To what extent have the evaluations results been used by Finnish stakeholders outside the MFA, e.g. Development Policy Committee, Parliament and CSOs?

RQ3: What are the typical factors facilitating or hampering the uptake of centralised evaluations?

3.1 What have been enabling factors for the use of evaluations and implementation of recommendations?

3.2 What have been the main reasons for not taking action on any given recommendation in the management response?

3.3 What have been the main reasons for not implementing agreed actions specified in the management response?

RQ4: How could the MFA better capture and retain knowledge and learning from evaluations as one tool for knowledge-based management?

7. Approach and methodology

The review draws from a broad and diverse body of information, evaluative evidence, and primary and secondary data. The main secondary data sources include the annual evaluation plans, evaluation reports, management responses and reports on the follow-up of management response.

Data from the different information sources are triangulated and analysed through content analysis, pattern matching and synthesis analysis. Human rights-based approach (HRBA), gender equality and non-discrimination are integrated into the review methodology as appropriate. Furthermore, ethical considerations are taken into account in the overall review approach.

The review team will finalize the methodology in the inception report.

8. Timing and reporting

The review will take place in 2022-2023. It will start in August 2022 by finalising the Terms of Reference and selecting the review team in September.

Inception phase: The inception phase starts right after the selection of the review team and ends with the approval of the inception report by 31 October 2022.

The (draft and final) inception report will include the review plan and initial desk study. The inception report will include the following sections: background and context; initial desk study findings (strategic level only), further development of the analytical framework; finalization of the methodology and summarized in a review matrix including review questions/sub-questions, judgment criteria, methods for data collection and analysis (the review matrix will reflect the use of a gender-sensitive

assessment framework and address relevant cross-cutting perspectives); final work plan and division of work between team members; tentative table of contents of the final report; possible data gaps; tentative implementation plan for stakeholder consultations with a clear division of work (participation, interview questions/guides/checklists, preliminary list of stakeholders and organizations to be contacted); communication and dissemination plan; analysis of risks and limitations and their mitigation; and budget. The structure of the review report and annexes will be agreed upon in the inception meeting.

Data collection and analysis: The data collection and analysis take place in November-December 2022 (please note that interviews should be done by latest in the beginning of December)

Briefings/FCR workshop: Stakeholder workshops (tentatively one for senior management and one for other relevant MFA staff and stakeholders) are held in mid-January 2023

Reporting and communication: The draft final report is expected to be submitted by 31 January 2023 and final report by 28 February 2023. The review results will be presented in March 2023.

Production of the first draft of the 4-pager for communication purposes will be the responsibility of the Consultant. The first draft will be provided simultaneously with the draft final report.

The final report will include a summary (including table on main findings, conclusions, and recommendations) in Finnish, Swedish, and English. The final report will be delivered in Word-format with all the tables and pictures also separately in their original formats. The revised reports have to be accompanied by a table of received comments and responses to them. In addition, the MFA requires access to the review's tools, data sets, or interim evidence documents, e.g., completed matrices, although it is not expected that these should be of publishable quality. The MFA treats these documents as confidential if needed.

Each deliverable is subject to specific approval. The review can move to the next phase only after receiving a written statement of acceptance by the MFA.

In addition to written deliverables, the review team is expected to participate in workshops and give oral presentations, often supported by PowerPoint slides. The review results will be presented in the MFA in Helsinki, with the review team members present.

The Consultant is expected to provide agreed visual materials, such as a minimum of 3 alternatives for the cover picture for EVA-11's acceptance.

9. Expertise required

The review team will consist of Team Leader and evaluator. The expertise requirements for the Team Leader/Team member are:

- For the Team Leader: Strong team leadership and management track record and commitment to delivering timely and high-quality evaluation reports;

- Fluency in Finnish;

- Strong familiarity with Finnish development policy and cooperation and related decision making procedures;

- Extensive evaluation experience in centralised, policy level evaluations in development policy and cooperation;

- Readiness to use a variety of review methods (e.g. participatory methods, in-depth interviews etc.) and hands-on experience in collecting and analysing quantitative and qualitative data;

- Readiness and availability to disseminate the review results and recommendations in the way that it supports learning of the MFA's staff and management;

- Good communication and people skills; ability to communicate with various stakeholders and to express ideas and concepts concisely and clearly in written and oral form;

- Should be flexible, available as well as able to commit and allocate sufficient amount of time to the entire evaluation process, including when faced with unexpected changes; and

- Expertise on applying and evaluating human rights-based approach in development policy and cooperation and enhancement of the rights of women and girls as well as persons with disabilities.

10. Quality assurance

The Consultant is expected to propose and implement a quality assurance system for the review. The proposal must specify the quality assurance process, methodology, tools and resources (QA personnel and resource allocations).

11. Budget

A total budget is estimated to be 55 000 Euros including a contingency for any unexpected expenses. The final budget will be decided during the Inception Phase. Travelling is not foreseen, and the travel costs are not thus included in the budget. Such expenses may be included later. Contingency will be used for those costs.

12. Mandate

The review team is entitled and expected to discuss matters relevant to this assessment with pertinent persons and organizations. However, it is not authorized to make any commitments on behalf of the Government of Finland or the Ministry. The assessment team does not represent the Ministry for Foreign Affairs of Finland in any capacity.

All intellectual property rights to the result of the Service referred to in the Contract will be the exclusive property of the Ministry, including the right to make modifications and hand over material to a third party. The Ministry may publish the result under Creative Commons license to promote openness and public use of assessment results.

13. Authorisation

Helsinki, 23.8.2022

Antero Klemola
Director
EVA-11
Ministry for Foreign Affairs of Finland

Annex 3. Approach, Methodology and Limitations

Review approach

The review of Use and Utility of Centralised Evaluations is based on a team and co-creation approach and has future orientation. Recent evidence in cognitive science provides the foundation for the review. Successful leaders understand organizational transformation that considers the physiological nature of the brain, and the ways in which it predisposes people to resist some forms of leadership and accept others can lead and influence mindful change.

Evaluations are conducted with the hope that identifying the source of the problem would help solve it. The ultimate purpose of this review is to see how information and knowledge that is produced in these centralised evaluations would be used in decision making, learning and accountability in the Ministry for Foreign Affairs. Our management models are based on the premise that knowledge is power. This "transmission" approach to exchanging information, exemplified by lectures and textbooks, where knowledge is "transmitted" to a passive receiver, has been the prevailing teaching method in academia.

Such a new way would be based on what the human brain wants in all the challenges of change, which is not knowledge but just to focus on solutions instead of problems, let people come to their own answers, and keep them focused on their insights. As Peter F. Drucker has said, "We now accept the fact that learning is a lifelong process of keeping abreast of change. And the most pressing task is to teach people how to learn."

Cognitive scientists are finding that people's mental maps, their theories, expectations, and attitudes, play a more central role in human perception than was previously understood. How, then, would we go about facilitating change? The impact of mental maps suggests that one way to start is by cultivating moments of insight. Large-scale behaviour change requires a large-scale change in mental maps. This in turn requires an event or experience that allows people to provoke themselves, in effect, to change their attitudes and expectations more quickly and dramatically than they normally would.

Focused attention to making organisational change like promoting learning organisation is not only an individual effort. Firstly, individual learning benefits that individual but not the wider organisation, and secondly it creates a lot of stress for a single individual change agent to push and pull the others in the unit or department. It is also difficult for human being to keep focused attention on many things at the same time. So focus is power when learning and change is prioritised to one or two things.

When the aim is to change something, there is always the need for human behaviour change, and that is more demanding than arranging materials, tools or a training event. The question is how we ensure that people start using the materials, adopt the new tools or act on new knowledge. This justifies the focus on learning theories adopted by the review team.

Thus, the methodology of the review has built on team approach in information collection and analysis. The idea was to share how to co-create learning and improving activities effectively and efficiently, and by that share a positive experiment of learning together and cater a learning culture in MFA, taking benefit of evaluations especially.

A draft Theory of Change (ToC) was drawn early on to identify the steps, actors and processes the review should focus on. It directed the review to look at processes in evaluation planning and management, people involved and organisational factors facilitating or hampering progress, and other actors. The EVA-11 was inspired by the Theory of Change approach, and with facilitation support from the review team developed a draft Theory of Change for the Unit itself (see Appendix 2). Since it is still considered an early draft, it has not been applied in the analysis.

Methodology

Inception phase

The approach and methodologies of the review were agreed with the EVA-11 during the inception in October 2022. Data collection and analysis tools were designed, and preliminary document review conducted. A review matrix that summarises the review methodology was also developed during the inception (see Appendix 1). In the matrix the main Review Questions are broken down to more detailed interview questions and discussion topics. Main sources of data per question are also there. Also the representatives of main stakeholder groups that were expected to be among the potential users of the evaluations were identified

Due to the tight schedule of the review, arrangements were made and invitations for planned co-sharing workshops with the Ministry people were issued already in October. The initial plan was to organise five co-sharing workshops for which approximately 40 people were expected to attend. However, there was little interest in this approach: initially less than 10 people accepted the invitation. Thus, in the inception meeting the plans were revised. Most of the planned co-creation workshops were substituted by individual discussions with MFA people.

Already during the inception it was mutually understood that documents provide valuable data about the evaluations, but they only speak of the past. The discussions with the MFA people and other stakeholders provide information about the present use and utility of evaluations and the potential for future which underlines the significance of using discussions as a source of information in the methodology.

Primary data – workshops, focus group discussions and individual discussions

The review team organised 36 different data collection sessions, mostly during November-December 2022. The sessions consisted of facilitated workshops, focus group discussions and individual or pair discussions. Several co-creation workshops were organised for MFA, Embassy and stakeholder groups (two for the MFA people, one for people from the Finnish Embassies, one for Finnish CSOs, and one for members of Development Policy Committee). Focus group discussions were organised for the Secretariat of Development Policy Committee, EMS coordinators and for consultants who have served as team leaders for centralised evaluations. The review team had three workshops with the team of the EVA-11. Otherwise the review team discussed with the MFA people individually. A few persons shared their views over email.

Altogether 62 persons shared their experiences and ideas with the review team. The persons consulted represent the MFA, the Finnish Embassies, the Development Policy Committee, the Parliament, Finnish Civil Society Organisations and private sector organisations.

Secondary data - documents

In this review documents are an important source of supplementary data. The documents studied consist of reports depicting different steps in the evaluation process from mandate of development evaluation to planning evaluations, managing and implementing evaluations, and finally addressing recommendations of the evaluations. The Development Evaluation Norm and Rules of Procedure of the Ministry are among key documents studied. Annual Evaluation Plans, Annual Reports of Development Evaluation, and the guidelines that the EVA-11 has prepared for decentralised evaluations were all studied. Reports of individual evaluations together with respective management responses and their follow-up reporting have been looked at. Other MFA documents related to development policy and development cooperation in the MFA were reviewed as well. The List of References is in Annex 4. The evaluation reports, their management responses and follow-up reports that were reviewed are listed in Annex 5.

Regarding the evaluation reports the focus has been analysing the recommendations, their clarity and the categories they fall in. With management responses the focus has been on the extent to which the recommendations are agreed (fully, partly, not at all) and identifying reasons why recommendations have not led into actions. With the follow-up reports the focus was on identifying if actions agreed in the follow-up table have actually been implemented or not and if there is any evidence of change resulting from the evaluation in the reports. The analysis of the management responses and the follow-up reports was made on the basis of the management response follow-up tables because they are the documents comparable to the evaluation recommendations.

Also, the annual evaluation work plans and annual evaluation reports produced by EVA-11 were analysed to have an overall picture of how centralised evaluations and evaluation function has evolved during 2015-2022.

Data analysis

Thus, the data used in the analysis consists of documents, and notes of workshops, focus group discussions and individual discussions. For most of the review questions documents provide only some/limited evidence – the workshops, focus group discussions and interviews have provided the key evidence for the review.

The review team analysed the data during the first half of January. In the analysis, each stream of evidence (discussion notes and documents) was systematically analysed against the Review Matrix questions. Data triangulation was a two-step process, i.e. qualitative discussion data was separately collated in an Excel file (as per the review sub-questions of the TOR). Key words were used to classify responses and used to identify narrative patterns of stakeholders views. The clustered data entries were then analysed against the questions in the Review Matrix and are the source of the findings based on primary data in the report. Views and opinions expressed by just one person were discarded from the analysis. Similarly, a systematic content analysis was conducted for the documents. The recommendations of the evaluation reports, the management response matrices and the follow-up reports were analysed to identify answers to relevant review questions.

Triangulation was applied both within primary data and between primary data and secondary data. As Chapter 3 shows, sometimes the views presented by the stakeholders agreed with each other and sometimes they contradict each other. Same applies to findings substantiated by the documents.

The analysis was boosted by the discussions with the MFA representatives in the FCR workshop in January 2023. The EVA-11 team together with the representatives from the Department for Africa and Middle East, Department for the Americas and Asia and Department for Development Policy participated in the workshop.

Reporting

The team wrote the 1st draft report in January, and it was submitted to the EVA-11 at the end of January 2023. The team revised the report extensively based on the comments received from the EVA-11 and the Final Draft Report was submitted in late February 2023. The Final Report will be completed by the end of March 2023.

Quality Assurance (QA) process

Quality assurance took place both during the inception and reporting phase. The QA expert has reviewed the draft Inception Report and the two draft reports. The team has addressed the comments made by the QA expert in the drafts before they were submitted to the MFA.

Limitations

Through the review process it was learned that most precious resource MFA staff are lacking at work is time. Organising information gathering through co-creation workshops proved to be challenging, and the review team had to alter plans. Although the planned duration of each workshop with MFA staff was only 1.5 hours and invitations were issued well ahead of time, only a handful of people expressed interest to participate. The opportunity for co-creation that the workshops would have provided was largely missed when individual discussions replaced most of the Ministry workshops.
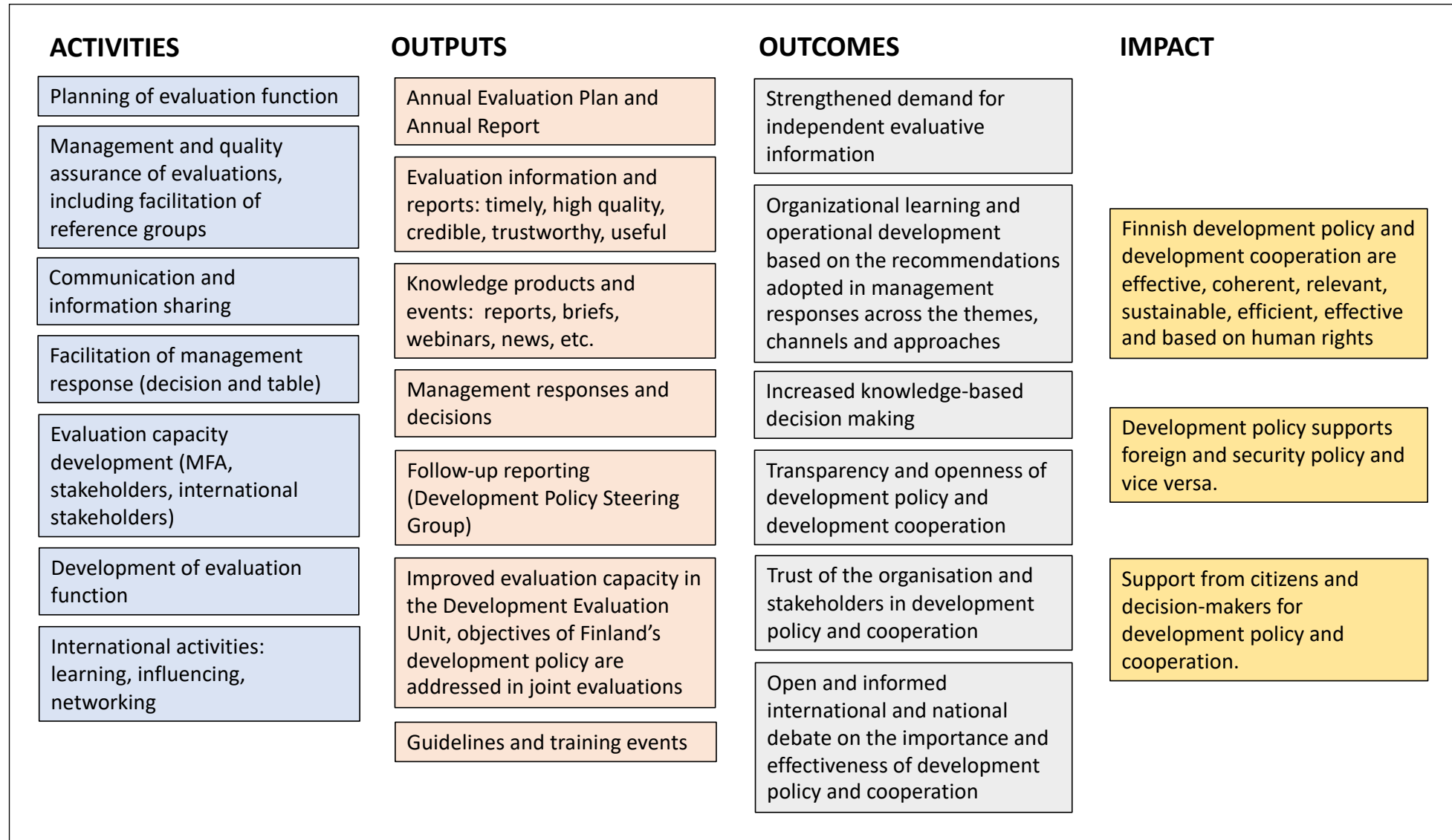
Appendix 1. The Review Matrix

REVIEW QUESTIONS	INTERVIEW QUESTIONS	KEY EVIDENCE (COLLECTED HOW / FROM WHERE)
RQ How transformative are centralised evaluations?		
1.1 To what extent centralised evaluations have influenced and/or changed policy dialogue inside and outside the MFA?	Please give examples, what centralised evaluation(s) have influenced your actions in policy dialogue? Where did you apply that new knowledge, approach etc.? How did you change your action, thinking or being? What was your learning of that situation? How does it affect your work today?	Workshops with MFA staff and CSO staff Focus Group Discussions (FGD) with Embassy staff and DPC members Key Informant Interviews (KII) with MFA staff, DPC Secretariat and DPC members (Parliamentarians)
1.2 To what extent centralised evaluations have influenced and/or changed policies, guidelines and processes in the MFA, including with regard to crosscutting objectives of Finnish development policy?	Please give examples, what centralised evaluation(s) have clearly influenced policies, guidelines and processes in the MFA? Who were the key stakeholders affected by this change? What was your reaction to the change? What worked well? Why did it work well? What was the change the evaluation result/ recommendation had in addressing the crosscutting objectives?	Workshops with MFA staff and Embassy staff KIIs with MFA staff, DPC Secretariat and DPC members (Parliamentarians) Document review
RQ2: How are centralised evaluations used and how useful they are?		
2.1 To what extent the evaluation processes (from planning until the follow up of the management response) support use of evaluation results and implementation of recommendations?	How well the current evaluation process is benefiting your work? What works well? What does not work at all? What little change could improve the current situation?	Workshops, FGDs and KIIs Document review (Evaluation plans, annual reports, management responses and their follow-up)
2.2 To what extent are the recommendations relevant, realistic and actionable?	What recommendations have been implemented in your unit? In your department? In your organisation? What is common with those recommendations that were not implemented?	Workshops, Interviews Selected reports, management responses
2.3 To what extent do the actions committed to in the management response adequately address the recommendations?	Have you participated in a working group for management response preparation? What is your experience of the discussion and collaboration? How transparently the management responses outline which recommendations will be endorsed and which not? What justifications are provided on the decisions? What change would further develop this collaboration?	KIIs of MFA staff Document review (management responses)

REVIEW QUESTIONS	INTERVIEW QUESTIONS	KEY EVIDENCE (COLLECTED HOW / FROM WHERE)
2.4 To what extent is there commitment to implement the recommended actions and how does that manifest itself at best?	From where do you get information on management responses? In what type of roles have you been in implementing activities agreed in the management response? How well did you succeed in addressing the recommendations adequately? What is your learning from these experiences? What needs to change so that you would benefit from the evaluations and improve your work?	KIIs of MFA staff
2.5 To what extent have the evaluations results been used by Finnish MFA, e.g. Development Policy Committee, Parliament and CSOs?	Which evaluation have you regarded useful for your work? To whom have you shared information of this evaluation? How did it affect discussion and decision making? What is the role of evaluations in steering development policy? Where should evaluation results be shared to reach wider public?	CSO workshop, KIIs of MFA staff, DPC and Parliamentarians
RQ3: What are the typical factors facilitating or hampering the uptake of centralised evaluations?		
3.1 What have been enabling factors for the use of evaluations and implementation of recommendations?	What factors make it easier to use evaluation results and implement recommendations? What factors encourage you? How are you inspiring others to take action?	KIIs of MFA staff Evaluation reports, e.g. Peer Review (2020) report
3.2 What have been the main reasons for not taking action on any given recommendation in the management response?	What factors impede or delay the implementation of agreed recommendations? What are the hidden rules of MFA organisational culture hampering the change? What is your tactic to hinder implementation of overburdening recommendations?	KIIs of MFA staff Workshops Document review (e.g. EVA 11 reports to Development Policy Steering Group, Annual Evaluation Reports)
3.3 What have been the main reasons for not implementing agreed actions specified in the management response?	Which type of actions agreed in the management response will not be implemented? Why have they not been implemented? Who could have taken the lead? What would make positive difference in the near future?	MFA Workshops, KIIs of MFA staff Document review, e.g. EVA-11 reports to Development Policy Steering Group and Annual Evaluation Reports
RQ4: How could the MFA better capture and retain knowledge and learning from evaluations as one tool for knowledge-based management?	What works well and provides you with knowledge and learning in the current evaluation process? What would strengthen these? What else is needed to better capture and retain knowledge and learning from evaluations? What should you / MFA let go?	Workshops (MFA, CSO), FGDs (Embassy, EMS team), KIIs of MFA staff

Appendix 2. Draft EVA-11 Theory of Change

ACTIVITIES

- Planning of evaluation function

- Management and quality assurance of evaluations, including facilitation of reference groups

- Communication and information sharing

- Facilitation of management response (decision and table)

- Evaluation capacity development (MFA, stakeholders, international stakeholders)

- Development of evaluation function

- International activities: learning, influencing, networking

OUTPUTS

- Annual Evaluation Plan and Annual Report

- Evaluation information and reports: timely, high quality, credible, trustworthy, useful

- Knowledge products and events: reports, briefs, webinars, news, etc.

- Management responses and decisions

- Follow-up reporting (Development Policy Steering Group)

- Improved evaluation capacity in the Development Evaluation Unit, objectives of Finland's development policy are addressed in joint evaluations

- Guidelines and training events

OUTCOMES

- Strengthened demand for independent evaluative information

- Organizational learning and operational development based on the recommendations adopted in management responses across the themes, channels and approaches

- Increased knowledge-based decision making

- Transparency and openness of development policy and development cooperation

- Trust of the organisation and stakeholders in development policy and cooperation

- Open and informed international and national debate on the importance and effectiveness of development policy and cooperation

IMPACT

- Finnish development policy and development cooperation are effective, coherent, relevant, sustainable, efficient, effective and based on human rights

- Development policy supports foreign and security policy and vice versa.

- Support from citizens and decision-makers for development policy and cooperation.

Annex 4. References

Berg A, Lähteenoja S, Ylönen M, Korhonen-Kurki K, Linko T, Lonkila K-M, Lyytimäki J, Salmivaara S, Salo H, Schönach P & Suutarinen I (2019). Path 2030 – An evaluation of Finland's sustainable development policy. Prime Minister's Office, 19.03.2019

Betts J, Mikkolainen P & Friedman J (2020). Evaluation of selected Finland's country strategies and country strategy approach for development cooperation with focus on fragile contexts. Volume 1 – Synthesis report. Ministry for Foreign Affairs, 2020.

Cooney G & Rojas K, Arsenault M & Babcock K (2015). Meta-evaluation of project and programme evaluations in 2012–2014. Ministry for Foreign Affairs, 2015.

de Man R, Vikman H, Tamrat I, Geheb, K & Toikka S (2021). Water as an Entrypoint for Peace Mediation. Evaluation on Finnish Water Diplomacy. Volume 1 – Main report. Ministry for Foreign Affairs.

Halme, K, Lähde K, Lamminkoski H, Mäkelä M & Giddings S (2019). Development Evaluation of Business with Impact (BEAM) Programme. Final Report. Ministry for Foreign Affairs.

Decree of the Ministry for Foreign Affairs on the Rules of Procedure of the Ministry for Foreign Affairs. 550/2008 (in Finnish: *Ulkoministeriön asetus ulkoministeriön työjärjestyksestä 28.8.2008/550*, available in: https://finlex.fi/fi/laki/ajantasa/2008/20080550 , accessed on 29 January 2023)

Finnish Government. 2020. Government Report on Finnish Foreign and Security Policy. Publications of the Finnish Government 2020:32.

Government Decree on the Ministry for Foreign Affairs 1171/2005 (in Finnish: *Valtioneuvoston asetus ulkoministeriöstä* 22.12.2005/1171), available in https://finlex.fi/fi/laki/ajantasa/2005/20051171 , accessed on 29 January 2023)

Karhu M & Lanki J (2022). In the Quest for True Commitment: The Finnish Aid Paradigm and the 0.7 Target in Domestic Politics. Forum for Development Studies, 49:3, 435-465, DOI: 10.1080/08039410.2022.2116353

Laaksonen S, Wallenius T, Huhta S, Nderitu A & Mikkolainen P (2021). Evaluation of Economic Development, Job Creation and Livelihoods. Final Report. Volume 1.1 – Main Report. Ministry for Foreign Affairs.

Laaksonen S, Palenberg M & Toikka S (2022). From Reactivity to Resilience. Assessment of the Response of Finnish Development Policy and Cooperation to the COVID-19 Pandemic. Ministry for Foreign Affairs.

LeBlanc R, Hennion M & Seppänen M (2016). Meta-evaluation of project and programme evaluations in 2014–2015. Ministry for Foreign Affairs, 2016.

Makkonen A (2021). Business to Government - a developmental review of the platform pilot (in Finnish). Ministry for Foreign Affairs.

MFA (2015a). Development evaluation norm of the Ministry for Foreign Affairs (MFA). Norm 1/2015 (unofficial translation accessed from https://um.fi/documents/35732/0/Development+evaluation+norm_MFA_2015+%281%29.pdf/dfdf4e72-a5b6-8628-d70b-a32f257f0c-2c?t=1624960477459). 21.01.2015. Ministry for Foreign Affairs.

MFA (2015b). Annual Evaluation Plan 2016–2018 (in Finnish: *Evaluoinnin kokonaissuunnitelma 2016-2018*). 9.12.2015. Ministry for Foreign Affairs, EVA-11

MFA (2016a). Annual Evaluation Plan 2017–2019 (in Finnish: Evaluoinnin kokonaissuunnitelma 2017-2019). 22.12.2016. Ministry for Foreign Affairs, EVA-11

MFA (2018a). 2017 Annual Report of Development Evaluation (In Finnish: *Kehitysevaluoinnin vuosiraportti 2017*). EVA-11, Ministry for Foreign Affairs.

MFA (2018b). Finland's Development Policy Results Report 2018. To Parliament. 1 November 2018. Ministry for Foreign Affairs.

MFA (2018c). Annual Evaluation Plan 2018-2020 (in Finnish: *Evaluoinnin kokonaissuunnitelma 2018-2020*). Ministry for Foreign Affairs, EVA-11

MFA (2018d). Reporting of evaluation management responses, HEL7M1595-8, 05.07.2018 (in Finnish: *Evaluointien johdon päätösten seurannasta raportointi*)

MFA (2019). Reporting of evaluation management responses, 5CG7213PN0-11, 02.10.2019 (in Finnish: *Evaluointien johdon päätösten seurannasta raportointi*)

MFA (2020a). Terms of Reference. Framework Agreement for Development Evaluation Services. Final 30/01/2020. Ministry for Foreign Affairs, EVA-11.

MFA (2020b). Guidelines for evaluation reference group. 16.11.2020. Ministry for Foreign Affairs, EVA-11.

MFA (2020c). Country Programme Template. Country Programme for Development Cooperation [name of the country] 2021–2024. Ministry for Foreign Affairs.

MFA (2020d). Internal instruction for influencing multilateral organisations (in Finnish: *Suomen avunantajaprofiilin vahvistaminen monenkeskisellä kehityspoliittisella vaikuttamisella: kokonaissuunnitelma ja työkalut kevään 2020 vuosikeskustelujen jälkeen*). 11.6.2020. Ministry for Foreign Affairs.

MFA (2020f). Annual Evaluation Plan 2020-2022 (in Finnish: Evaluoinnin kokonaissuunnitelma 2020-2022). 22.01.2020. Ministry for Foreign Affairs, EVA-11

MFA (2020g). Annual Evaluation Plan 2021-2023 (in Finnish: Evaluoinnin kokonaissuunnitelma 2020-2022). 21.12.2020. Ministry for Foreign Affairs, EVA-11

MFA (2021a). Report on Development Policy Extending Across Parliamentary Terms. Ministry for Foreign Affairs. http://urn.fi/URN:ISBN:978-952-383-709-6

MFA (2021b). The long path of learning from evaluation recommendation to effective implementation (in Finnish: *Oppimisen pitkä polku evaluointien suosituksista niiden vaikuttavaan toimeenpanoon.*) Draft 17.3.2021. Ministry for Foreign Affairs.

MFA (2021c). Toolkit for management response working groups. May 2021. Ministry for Foreign Affairs, EVA-11.

MFA (2021d). Annual Evaluation Plan 2022–2024 in Finnish: Evaluoinnin kokonaissuunnitelma 2022-2024). Ministry for Foreign Affairs, EVA-11

MFA (2021f). A review of Finland's cooperation supporting peace processes (in Finnish: *Katsaus Suomen ulkoministeriön rauhanprosesseja tukevaan yhteistyöhön*). Ministry for Foreign Affairs, EVA-11

MFA (2022a) Finland's development aid appropriations and disbursements 1989-2021, available in https://um.fi/development-cooperation-appropriations (accessed on 8 December 2022)

MFA (2022b) Finland's Development Policy Results Report 2022. To Parliament. 9 November 2022 (in Finnish). Ministry for Foreign Affairs.

MFA (2022c). International Instruction for Development Cooperation Quality Group (In Finnish: *Kehitysyhteistyön laaturyhmäohje, päivitetty versio 09.05.2022*). Ministry for Foreign Affairs.

MFA (2023). Evaluation Manual. In: https://um.fi/development-cooperation-evaluation-manual, accessed on 29 January 2023.

Ministry of Finance (2023). Evidence-based management. In: https://vm.fi/en/evidence-based-management, accessed on 28 January 2023.

Mintzberg H, Etzion D & Mantere S (2018). Worldly Strategy for the Global Climate. In Stanford Social Innovation Review, 16(4): 42-47

National Audit Office (2021). Finland's International climate finance – Steering and Effectiveness. Performance audit report. Translation of performance audit report 6/2021.

OECD (2023). Accountable and Effective Development Co-operation. 2022 Review of Evaluation Systems in Development and Humanitarian Assistance. Draft. OECD 2023

OECD/DAC (2019). Better Criteria for Better Evaluation. Revised Evaluation Criteria Definitions and Principles of Use. OECD/DAC Network on Development Evaluation, December 2019. In: https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm

OECD/DAC (2023). OECD/DAC (2023) Official Development Assistance (ODA), in: https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/official-development-assistance.htm, accessed on 27 January 2023

Palenberg M, Bartholomew A, Mayne J, Mäkelä M & Esche L (2019). How do we learn, Manage, and Make Decisions in Finland's Development Policy and Cooperation. Management of results information and knowledge at the Ministry for Foreign Affairs of Finland. Final Report. Ministry for Foreign Affairs of Finland.

Peer Review of Four Evaluation Functions (2020). EVA-11, Ministry for Foreign Affairs, Finland. Evaluation and Audit Unit, Department of Foreign Affairs and Trade, Ireland. Evaluation and Corporate Controlling Division, Agency for Development Cooperation, Ministry of Foreign Affairs, Switzerland. Evaluation Unit, Quality and Resources Section, Economic Cooperation and Development Division, State Secretariat for Economic Affairs, Switzerland. November 2020

Presencing Institute - Otto Scharmer (2023). Theory of U. In: https://www.u-school.org/theory-u , accessed on 20 January 2023

Programme of Prime Minister Sanna Marin's Government (2019). Inclusive and competent Finland - a socially, economically and ecologically sustainable society. Publications of the Finnish Government 2019:3, 10 December 2019. https://valtioneuvosto.fi/en/marin/government-programme

Rassmann K, Byron G, Poutiainen P & Mikkolainen P (2018). Evaluation on improvement of women's and girls' rights in Finland's development policy and cooperation. Ministry for Foreign Affairs, 2018.

Senge P (1990). The Fifth Discipline.

Senge PM, Kleiner A, Roberts C, Roth G & Ross R (1999). The Dance of Change. The Challenges of Sustaining Momentum in Learning Organisations. Doubleday (Random House, Inc.), 1999.

Silvestrini S, Väth SJ, Römling C, Lieckefett M & Mikkolainen P (2018) Meta-evaluation of project and programme evaluations in 2015–2017. Ministry for Foreign Affairs, 2016.

Väth J, Silvestrini S, Gaus H, Mikkolainen P, Flaig M & Wicke J (2022). Meta-evaluation of project and programme evaluations in 2017–2020. Ministry for Foreign Affairs, 2022.

Annex 5. Evaluation reports, management matrices and follow-up reports used in the review

The review has used the following evaluation reports (23), management responses (20) and follow-up reports of management responses (12) in the analysis of recommendations, management responses and follow-up reports.

Table 1. List of evaluation reports, management responses and follow-up reports used in the review

NAME OF EVALUATION AND YEAR PUBLISHED	MATERIALS USED IN ASSESSING RQ2 AND RQ3:		
	EVALUATION REPORT	MANAGEMENT RESPONSE	FOLLOW-UP REPORT
Evaluation of Finland's Humanitarian Assistance 2016-2022 (2022)	X	-	-
Evaluation of the Finnish Development Policy influencing in the European Union (2022)	X	-	-
Meta-evaluation of project and programme evaluations in 2017-2020 (2022)	X	X	-
Evaluation on the Transition Process of Finnish Vietnamese Cooperation in 2008-2020	X	X	-
Evaluation on Development Cooperation Carried out by the Department of Russia, Eastern Europe and Central Asia, including the Wider Europe Initiative (WEI) (2021)	X	X	-
Water as an Entry Point for Piece Mediation. Evaluation on Finnish Water Diplomacy (2021)	X	X	-
Evaluation of Economic Development, Job Creation and Livelihoods (2021)	X	X	-
Evaluation of Finnish Development Policy Activities in Multilateral Organisations (2020)	X	X	-
Adapting for Change: Country Strategy Approach in Fragile Contexts (2020)	X	X	X (interim)
"How do we Learn, Manage and Make Decisions in Finland's Development Policy and Cooperation (2019)	X	X	-
Evaluation on Forced Displacement and Finnish Development Policy (2019)	X	X	X
Developmental Evaluation of Business with Impact 2015-2019 (BEAM) Programme (2019)	X	X	X

NAME OF EVALUATION AND YEAR PUBLISHED	MATERIALS USED IN ASSESSING RQ2 AND RQ3:		
	EVALUATION REPORT	MANAGEMENT RESPONSE	FOLLOW-UP REPORT
Meta-evaluation of project and programme evaluations in 2015-2017 (2018)	X	X	X
Evaluation on Improvement of Women's and Girls rights (2018)	X	X	X
Meta-evaluation of project programme evaluations 2014-2015 (2017)	X	(X)	(X)
Programme-based Support Through Finnish Civil Society Organizations (Meta analysis 2017)	X	X	X
Evaluation on Finland's Development Cooperation Country Strategies and Country Strategy Modality (2016)	X	X	X
Evaluation of Aid for Trade (2016)	X	X	X
Evaluation of Humanitarian Mine Action (2015)	X	X	X
Evaluation of Inclusive Education in Finland's Development Cooperation 2004-2013 (2015)	X	X	X
Evaluation on Finland's Development Policy Programmes from Results-Based Management Point of View (2015)	X	X	X
Meta-evaluation of project and programme evaluations in 2012-2014 (2015)	X	X	X
Evaluation on Finland's Development Cooperation with Kenya 2007-2013 (2015)	X	X	X

Annex 6. Evaluation Process

In this Annex, the evaluation process that the EVA-11 applies for all centralised evaluations is described.

The EVA-11 prepares annually **a comprehensive evaluation plan** for the MFA. The document presents a detailed plan for both centralised and decentralised evaluations that are planned to be initiated within the year. A pipeline for evaluations for years 2 and 3 is also included in the plan. The annual evaluation plan is discussed in the Development Policy Steering Group. Subsequently, the Under-Secretary of State responsible for development policy confirms the plan.

MFA commissions all evaluations and reviews to independent evaluators and expert teams. The EVA-11 started using **framework contracts for the centralised evaluations** in 2017 (MFA 2018a). According to the Terms of Reference of the current Evaluation Management Services (EMS) framework contract (MFA 2020a), the EVA-11 is a commissioner of evaluations and is responsible for planning of the evaluation, leading the design and conducting of evaluations, quality assurance of evaluation deliverables, dissemination and use of the evaluations. The EMS Service Provider is responsible for providing headhunting services for the evaluation expertise needed in evaluation teams and coordinates and conducts the evaluations.

In the EMS model (MFA 2020a), each individual evaluation consists of four main steps, namely evaluation planning and preparations, conducting evaluations, report writing and publishing, and management response and follow-up. For each step, roles and responsibilities have been defined for the EVA-11 and the Service Provider. The main responsibilities of the EVA-11 are listed in Table 1.

Table 1. Main tasks of the EVA-11 during an evaluation applying the EMS model

STEP	MAIN TASKS OF THE EVA-11
Evaluation planning and preparations	Preparing the Concept Note for the evaluation where background, tentative purpose, objectives, scope and use of evaluation are defined; Approving the Team Leader and team members; and Approving the Terms of Reference for the evaluation
Conducting evaluations	Approving all deliverables during the evaluation (reports, communication materials, presentations); and Appointing and organising reference group meetings
Report writing and publishing	Organising, collecting and compiling feedback of the relevant stakeholders; Organising dissemination seminars and webinars; Publishing and disseminating the reports.
Management response and follow-up	Facilitating the management response; Facilitating the follow-up

Source: MFA (2020a). Terms of Reference. Framework Agreement for Development Evaluation Services. Final 30/01/2020. Ministry for Foreign Affairs, EVA-11.

For each evaluation, **a reference group** is established. The purpose of reference group is to support a useful, credible, transparent, impartial, and quality evaluation process. It is expected that participation and engagement of stakeholders would lead to better quality, utility, ownership, and use of evaluation. Reference groups are also expected to contribute to improved organisational learning. The EVA-11 chairs the reference group and asks relevant MFA units, departments and/or Embassies to assign representative(s) to the group. Also external stakeholders, e.g. civil society partners, other Finnish agencies, academia, partner countries and donors can be members of a reference group (MFA 2020b).

The evaluation norm stipulates that "*all development evaluations carried out in the MFA are responded to by **a management response**". The EVA-11 initiates the preparation of the management response and acts as a secretary of the group. Otherwise, the responsibility of preparing the response is in the hands of relevant departments and units. The draft management response prepared by the group is discussed at the Development Policy Steering Group and the Under-Secretary of State responsible for development policy signs it (MFA 2015a). During the period of review, the management response has consisted of two documents, a management response decision (published in the MFA website), and a follow-up table with information about agreed actions (available upon request). The EVA-11 has produced a guideline and templates to support preparation of the management response (MFA 2021c). The management response working group has the responsibility to decide which evaluation recommendations they consider relevant and propose to be approved and which recommendations should not be approved. The group will then identify activities to implement the approved recommendations and agree about the responsibilities and the schedule for implementing the activities.

Follow-up reporting of evaluations takes place two years after the management response decision has been signed. The EVA-11 assists the departments responsible for implementing the agreed management response actions in monitoring the progress. The Unit prepares periodical summaries and lessons learned for the Development Policy Steering Group. The Unit also coordinates the preparation of the follow-up report (MFA 2015a, MFA 2021c).

Annex 7. Learning theories

As learning from evaluations is the main focus of this review, and as there are plenty of research of the topic, theories of three academics are introduced due to their relevance to this review.

Peter Senge (Massachusetts Institute of Technology (MIT, USA) laid down the foundation for learning in organisations already in the 1990s. Henry Mintzberg (McGill University, Canada) has focused on how organisations operate, and Otto Scharmer (MIT) is co-leading a new paradigm for transforming organisations and societies to the sustained future. Among other things, Scharmer is a member of the UN Learning Advisory Council for the 2030 Agenda.

Their work and models have inspired the review team and provide the review with frameworks to create a bigger picture and understanding of where to focus when the aim is to foster learning culture, navigate change and thrive as individuals, teams, organisations and as a society.

The review has focused on learning theories because so many corrective activities in the past have emphasised training and information management systems in MFA. When the aim is to change something, there is always the need for human behaviour change, and that is more demanding than arranging materials, tools, or a training event. The question is how do we ensure that people start using the materials, adopt the new tools or act on new knowledge?

The answer is to see learning as inseparable from everyday work. **The most important learning occurs in the context of our daily life**, the aspirations we pursue, the challenges we face, and the responses we bring forth – individually and collectively. All organisations learn – in the sense of adapting as the world around them changes. But some organisations are faster and more effective learners (Senge et al. 1999). So it is about learning whenever it is talk about change in thinking, being or action.

How to become more effective in learning and more effective in changing behaviour at organisational context? According to Senge (1999) **five learning disciplines are fundamental for organisation wide learning**. Individual learning is about cultivating **personal mastery**, and **self-reflection** on own attitudes and perceptions that influence thought and interaction. **Shared vision** is necessary both to nourish commitment and to develop shared images of the future. **Team learning** through dialogue draws forth intelligence and ability greater than the sum individual's talent. Last but not least learning of interdependencies, that is **systems thinking** increases understanding of forces that facilitate or hamper the intended changes.

What we pay attention to, and how we pay attention - both individually and collectively - is key to what we create. How to activate our capacity to lean into the emerging future may well be the most important **leadership challenge of our time**. How do you cultivate curiosity, compassion, and courage in the face of prejudice, anger, and fear? The new school of transformation, Theory U., co-created at MIT, provides us with an overall theory how to lead such change in businesses, government organisations and civil society organisations worldwide (Presensing Institute – Otto Scharmer 2023).

Theory U is an awareness-based method for changing systems. It blends systems thinking, innovation, and leading change - from the viewpoint of an evolving human consciousness. It integrates organizational learning in the tradition of Peter Senge.

One head is not enough in this rapid change, and more team relationships are needed. Stakeholder relationships need to be changed from **transactional to transformative**. It is all about co-creation and leading from the emerging future.

Theory U offers a method for rethinking the parts and the whole by making it possible for the system to sense and see itself. When that happens, the collective consciousness begins to shift from ego-system awareness to eco-system awareness - from a silo view to a systems view. At its core, Theory U comprises three main elements:

1. A framework for seeing the blind spot of leadership and systems change: the "interior condition" from which we operate;

2. A method for implementing awareness-based change. The focus is on building the collective capacity to shift the inner place from which we operate;

3. A new narrative for evolutionary societal change: updating our mental and institutional operating systems in all of society's sectors. What does it take to redesign societies in ways that address the pressing challenges of our time? What does it take to apply the power of mindfulness and awareness to the transformation of the collective system?

In a rapidly changing world, creating the future involves rapid prototyping rather than long-term planning. Effective leadership involves the capabilities of sense-making, developing relationships, visioning, and inventing new ways of organizing. To make social responsibility work, leaders need to develop the skill of facilitating collaboration among key actors from all three sectors: business, government, and civil society. You cannot solve today's leadership problems with the same level of consciousness that created them (paraphrasing Einstein). The theory U. is a transformational approach, providing tools and methods for action-oriented learning.

Mintzberg is known for his profound research on organisations and strategy, strategy in the sense of how organisations operate (Mintzberg et al 2018). He promotes the idea of strategy as learning and adaptation, and that intentions and planning play a crucial role as well. For Mintzberg, an organisation's actual strategy (its "realised strategy") is always a mix of what it planned (its "intended strategy") and the unexpected events to which it had to respond ("emergent strategy"). Learning as strategic option can also face problems and risks. The following three situations make big risks for organisation success:

- If there is no intended strategy at all then work gets fragmented.

- The strategy is lost, if the valid and useful strategy is given up, and working with new things is started only because they are new and interesting.

- The strategy is wrong, if the organisation is drifting through small steps and ostensibly small decisions into an unfavourable position.

The review understands that the MFA is a bureaucracy, in which departmental governance, political rules, regulations and conflicting pressures, sudden public events, annual budgets and private-public interdependencies guide decision making and learning.

Annex 8. Finnish ODA disbursements 2013-2021

Table 1. Finnish ODA disbursements 2013-2021, MEUR – Exclusive ODA and other ODA

ITEM / BUDGET LINE	DISBURSEMENTS, MEUR/YEAR								
	YEAR								
EXCLUSIVE ODA, BUDGET ITEM/LINE	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**
Multilateral development cooperation	278.4	351.7	343.9	141.7	163.9	170.1	171.4	233.4	244.0
Country-specific and regional development cooperation	247.3	290.0	254.6	202.6	153.6	118.2	140.7	165.6	157.6
European Development Fund (EDF)	47.1	47.8	50.0	52.8	59.7	67.8	70.8	70.8	57.1
Non-country specific development cooperation	60.7	61.2	48.8	38.4	31.1	32.1	31.0	38.8	54.2
Humanitarian assistance	96.4	105.7	97.8	84.0	73.3	72.1	78.1	113.4	104.5
Planning, support functions and communication	10.6	8.0	7.9	5.7	4.9	4.0	4.4	5.1	4.6
Evaluation and internal audit	2.2	1.9	2.7	3.0	2.4	1.3	1.8	2.0	2.1
Support conducted by civil society organisations	105.1	109.9	113.3	69.6	67.5	63.7	63.5	75.1	78.9
Concessional credits	14.2	15.1	7.6	7.4	8.7	7.7	6.8	4.1	7.9
Exclusive ODA budget items, total	**861.9**	**991.3**	**926.6**	**605.2**	**565.0**	**536.9**	**568.5**	**708.4**	**710.7**
Other ODA disbursements									
Finnfund	20.2	39.6	26.2	22.5	27.7	43.6	52.8	79.0	44.7
Financial instruments (excl. Finnfund)	0.0	0.0	0	9.2	68.0	0.0	96.2	20.0	130.7
Finland's share of EU cooperation budget	102.3	100.3	104.2	142.6	162.0	135.5	140.4	162.9	184.1
Administration costs	53.0	55.0	36.5	31.0	32.3	29.6	30.4	29.3	31.2
Refugee reception costs	15.7	12.1	35.2	117.7	68.7	48.3	80.1	51.2	58.5
Civilian crisis management costs	15.5	18.4	16.7	17.3	15.6	15.6	13.4	15.7	11.8
Other core support to multilateral organisations	10.5	13.8	14.4	10.5	21.9	22.4	26.1	47.0	36.4
Other items counted as ODA	2.0	1.5	1.2	0.3	0.2	1.9	2.1	7.9	6.0
Other ODA disbursements, total	**219.2**	**240.7**	**234.5**	**351.0**	**396.4**	**296.9**	**441.6**	**412.9**	**503.4**
Total ODA disbursements	**1081.1**	**1232.0**	**1161.1**	**956.2**	**961.4**	**833.7**	**1010.1**	**1121.4**	**1214.1**

Source: MFA (2022a) Finland's development aid appropriations and disbursements 1989-2021, available in https://um.fi/development-cooperation-appropriations (accessed on 8 December 2022)

REVIEW OF THE USE AND UTILITY OF CENTRALISED EVALUATIONS • FINAL REPORT